

C10

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
27 December 2001 (27.12.2001)

PCT

(10) International Publication Number
WO 01/98456 A2

- (51) International Patent Classification⁷: C12N
- (21) International Application Number: PCT/US01/19157
- (22) International Filing Date: 15 June 2001 (15.06.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/212,726 20 June 2000 (20.06.2000) US
- (71) Applicants (for all designated States except US): GENE LOGIC, INC. [US/US]; 708 Quince Orchard Road, Gaithersburg, MD 20878 (US). JAPAN TOBACCO, INC. [JP/JP]; 2-1, Toranomom 2-chome, Manato-ku, Tokyo 105-8422 (JP).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): MUNGER,

William, E. [US/US]; 708 Quince Orchard Road, Gaithersburg, MD 20878 (US). KULKARNI, Prakash [US/US]; 708 Quince Orchard Road, Gaithersburg, MD 20878 (US). GETZENBERG, Robert, H. [US/US]; 708 Quince Orchard Road, Gaithersburg, MD 20878 (US).

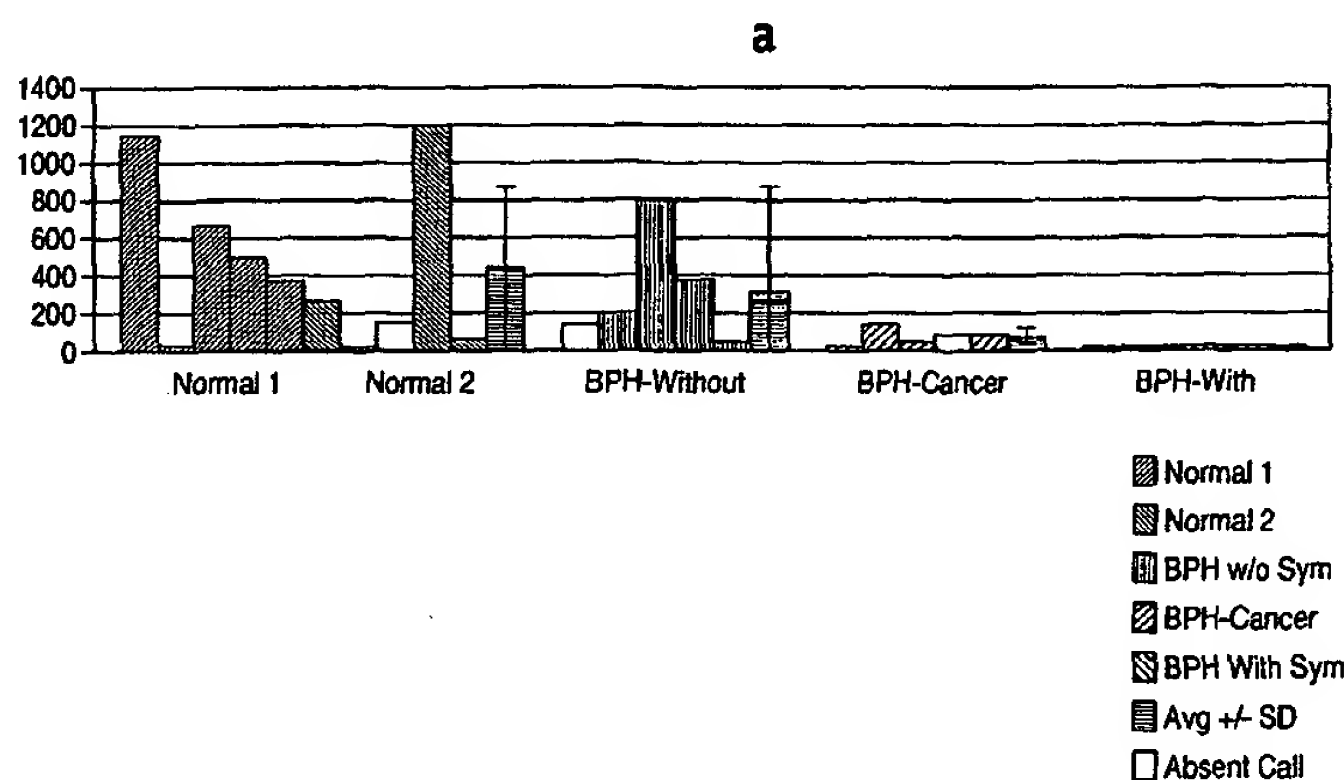
(74) Agent: TUSCAN, Michael, S.; Morgan, Lewis & Bockius LLP, 1800 M Street, N.W., Washington, DC 20036-5869 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

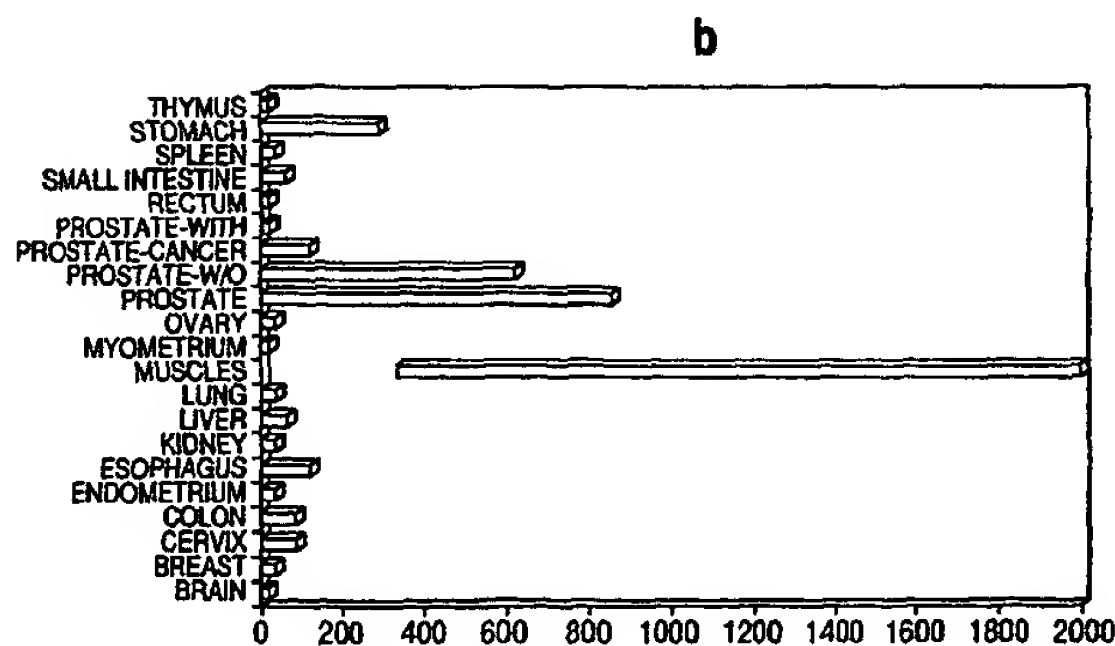
(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European

[Continued on next page]

(54) Title: IDENTIFICATION OF cDNAs ASSOCIATED WITH BENIGN PROSTATIC HYPERPLASIA



(57) Abstract: The invention relates generally to the changes in gene expression in Benign Prostatic Hyperplasia (BPH). The invention relates specifically to two novel human genes which correspond to mRNA species that are differentially expressed in BPH compared to normal prostate tissue.



WO 01/98456 A2



patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

— *entirely in electronic form (except for this front page) and available upon request from the International Bureau*

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

IDENTIFICATION OF cDNAs ASSOCIATED WITH BENIGN PROSTATIC HYPERPLASIA

INVENTORS

William E. Munger, Prakash Kulkarni and Robert H. Getzenberg

RELATED APPLICATIONS

This application claims priority to U.S. Provisional Application No. 60/212,726, filed June 20, 2000 and is herein incorporated by reference in its entirety.

FIELD OF THE INVENTION

5 The invention relates generally to the changes in gene expression in prostate tissue removed from male patients with benign prostatic hyperplasia (BPH). The invention specifically relates to two novel human genes which are differentially expressed in BPH tissue compared to normal prostate tissue.

10 BACKGROUND OF THE INVENTION

 BPH is the most common benign tumor in men aged > 60 years. It is estimated that one in four men living to the age of 80 will require treatment for this disease. BPH is usually noted clinically after the age of 50, the incidence increasing with age, but as many as two thirds of men between the ages of 40 and 49 demonstrate histological evidence of the
15 disease.

 The anatomic location of the prostate at the bladder neck enveloping the urethra plays an important role in the pathology of BPH, including bladder outlet obstruction. Two prostate components are thought to play a role in bladder outlet obstruction. The first is the relative increased prostate tissue mass. The second component is the prostatic smooth
20 muscle tone.

 The causative factors of BPH in man has been intensively studied. See Ziada *et al.*, (1999) *Urology* 53:1-6. In general, the two most important factors appear to be aging and the presence of functional testes. Although these factors appear to be key to the development of BPH, both appear to be nonspecific.

Molecular Changes in BPH

Little is known about the molecular changes in prostate cells associated with the development and progression of BPH. It has been demonstrated that the expression levels of a number of individual genes are changed compared to normal prostate cells. These changes in gene expression include a decreased level of Wilm's tumor gene (WT-1) and increased expression of insulin growth factor II (IGF-II) (Dong *et al.*, (1997) *J Clin Endocrin Metab* 82:2198-2203).

While the changes in the expression levels of a number of individual genes have been identified, the investigation of the global changes in gene expression has not been reported. Accordingly, there exists a need for the investigation of the changes in global gene expression levels as well as the need for the identification of new molecular markers associated with the development and progression of BPH. Furthermore, if intervention is expected to be successful in halting or slowing down BPH, means of accurately assessing the early manifestations of BPH need to be established. One way to accurately assess the early manifestations of BPH is to identify markers which are uniquely associated with disease progression. Likewise, the development of therapeutics to prevent or stop the progression of BPH relies on the identification of genes responsible for BPH growth and function.

SUMMARY OF THE INVENTION

The present invention is based on the discovery of two new genes that are differentially expressed in BPH tissue compared to normal prostate tissue. The invention includes isolated nucleic acid molecules selected from the group consisting of an isolated nucleic acid molecule that encodes the human amino acid sequence of SEQ ID NO: 2 or 4, an isolated nucleic acid molecule that encodes a fragment of at least 6 amino acids of SEQ ID NO: 2 or 4, an isolated nucleic acid molecule which hybridizes to a nucleic acid molecule comprising SEQ ID NO: 1 or 3 and an isolated nucleic acid molecule which hybridizes to a nucleic acid molecule that encodes the amino acid sequence of SEQ ID NO: 2 or 4. Nucleic acid molecules of the invention may have about 60% nucleotide sequence identity through the open reading frame of SEQ ID NO: 1 or 3, preferably about 70-75% sequence identity, more preferably about 80-85% sequence identity, and even more preferably at least about 90% sequence identity through the open reading frame.

The present invention further includes the nucleic acid molecules operably linked to one or more expression control elements, including vectors comprising the isolated nucleic acid molecules. The invention further includes host cells transformed to contain the nucleic acid molecules of the invention and methods for producing a protein comprising the step of
5 culturing a host cell transformed with a nucleic acid molecule of the invention under conditions in which the protein is expressed.

The invention further provides an isolated polypeptide selected from the group consisting of an isolated polypeptide comprising the amino acid sequence of SEQ ID NO: 2 or 4, an isolated polypeptide comprising a functional fragment of at least 10 amino acids of
10 SEQ ID NO: 2 or 4, an isolated polypeptide comprising conservative amino acid substitutions of SEQ ID NO: 2 or 4 and naturally occurring amino acid sequence variants of SEQ ID NO: 2 or 4. Polypeptides of the invention also include polypeptides with an amino acid sequence having at least about 50%, 60%, 70% or 75% amino acid sequence identity with the sequence set forth in SEQ ID NO: 2 or 4 more preferably at least about 80%, even
15 more preferably at least about 90%, and most preferably at least about 95% sequence identity with the sequence set forth in SEQ ID NO: 2 or 4.

The invention further provides an isolated antibody that specifically binds to a polypeptide of the invention, including monoclonal and polyclonal antibodies.

The invention further provides methods of identifying an agent which modulates the
20 expression of a nucleic acid encoding a protein of the invention, comprising the steps of: exposing cells which express the nucleic acid to the agent; and determining whether the agent modulates expression of said nucleic acid, thereby identifying an agent which modulates the expression of a nucleic acid encoding the protein.

The invention further provides methods of identifying an agent which modulates at
25 least one activity of a protein of the invention, comprising the steps of: exposing cells which express the protein to the agent; and determining whether the agent modulates at least one activity of said protein, thereby identifying an agent which modulates at least one activity of the protein.

The invention further provides methods of identifying binding partners for a protein
30 of the invention, comprising the steps of: exposing said protein to a potential binding partner; and determining if the potential binding partner binds to said protein, thereby identifying binding partners for the protein.

The present invention further provides methods of modulating the expression of a nucleic acid encoding a protein of the invention, comprising the step of: administering an effective amount of an agent which modulates the expression of a nucleic acid encoding the protein. The invention also provides methods of modulating at least one activity of a protein of the invention, comprising the step of: administering an effective amount of an agent which modulates at least one activity of the protein.

The present invention further includes non-human transgenic animals modified to contain the nucleic acid molecules of the invention or mutated nucleic acid molecules such that expression of the polypeptides of the invention is prevented.

The invention further provides methods of diagnosing BPH or other disease states, comprising: determining the level of expression of a nucleic acid molecule of the invention or polypeptide of the invention.

The invention further includes compositions comprising a diluent and a polypeptide or protein selected from the group consisting of an isolated polypeptide comprising the amino acid sequence of SEQ ID NO: 2 or 4, an isolated polypeptide comprising a fragment of at least 6 amino acids of SEQ ID NO: 2 or 4, an isolated polypeptide comprising conservative amino acid substitutions of SEQ ID NO: 2 or 4, naturally occurring amino acid sequence variants of SEQ ID NO: 2 or 4 and an isolated polypeptide with an amino acid sequence having at least about 35%, 40%, 50%, 60%, 70% or 75% amino acid sequence identity with the sequence set forth in SEQ ID NO: 2 or 4, preferably at least about 80%, more preferably at least about 90%, and most preferably at least about 95% sequence identity with the sequence set forth in SEQ ID NO: 2 or 4.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 - The top part of Figure 1 shows the expression of JT460914 across normal human control and BPH samples, including BPH samples from men without symptoms, BPH samples from men without symptoms who were diagnosed with prostate cancer, and BPH samples from men with symptoms. In all cases, the subregion of the prostate analyzed was the transitional zone. The lower part of Figure 1 shows an electronic Northern, in which the expression level of JT460914 was measured across a panel of normal tissues using the Affymetrix 42K human GeneChip set. For each tissue type, the mean +/- SDM is shown as a horizontal bar graph for samples obtained from 3 or more normal individuals.

Figure 2 - Figure 2 shows the results of semi-quantitative PCR for the expression of JT460914 in various tissues. The samples in each lane are as follows: M) low mass DNA markers; 1) heart; 2) brain; 3) leukocytes; 4) lung; 5) liver; 6) fetal brain; 7) kidney; 8) spleen; 9) placenta; 10) prostate from BPH patient; 11) glomeruli; 12) osteoblasts. The samples in the left panels were amplified by 25 cycles of PCR; the samples in the right panels by 30 cycles. The primers used with the samples in the upper panels were F15-36 and R323-344, while the primers used with the samples in the lower panels were F40-61 and R323-344.

Figure 3 - Figure 3 is a PEPLOT of JT460914.

Figure 4 - The top part of Figure 4 shows the expression of JT156897 across normal human control and BPH samples, including BPH samples from men without symptoms, BPH samples from men without symptoms who were diagnosed with prostate cancer, and BPH samples from men with symptoms. In all cases, the subregion of the prostate analyzed was the transitional zone. BPH with symptoms is defined primarily as the need for frequent urination, especially nocturnally. The lower part of Figure 4 shows an electronic Northern, in which the expression level of JT156897 was measured across a panel of normal tissues using the Affymetrix 42K human GeneChip set. For each tissue type, the mean +/- SDM is shown as a horizontal bar graph for samples obtained from 3 or more normal individuals.

Figure 5 - Figure 5 shows the results of semi-quantitative PCR for the expression of JT156897 in various tissues. The samples in each lane are as follows: M) low mass DNA markers; 1) heart; 2) brain; 3) leukocytes; 4) lung; 5) liver; 6) fetal brain; 7) kidney; 8) spleen; 9) placenta; 10) prostate from BPH patient; 11) glomeruli; 12) osteoblasts. The samples in the left panel were amplified by 25 cycles of PCR; the samples in the right panel by 30 cycles. The primers used with these samples were F8 and R258.

Figure 6 - Figure 6 is a PEPLOT of JT156897.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

I. General Description

The present invention is based in part on the identification of two new gene families that are differentially expressed in human BPH tissue compared to normal human prostate tissue. These gene families correspond to two human genes (SEQ ID NOS: 1 and 3) that encode proteins of 329 (JT460914) and 642 (JT156897) amino acids (SEQ ID NOS: 2 and

4, respectively). Genes that encode these human proteins may also be found in other animal species, particularly mammalian species.

The proteins can serve as diagnostic markers or agents and as targets for agents that modulate gene expression or activity. For example, agents may be identified that modulate biological processes associated with prostate growth, including the hyperplastic process of BPH.

The present invention is further based on the development of methods for isolating binding partners that bind to the proteins. Probes based on the proteins are used as capture probes to isolate potential binding partners, such as other proteins. Dominant negative proteins, DNAs encoding these proteins, antibodies to these proteins, peptide fragments of these proteins or mimics of these proteins may be introduced into cells to affect function. Additionally, the proteins provide novel targets for the screening of synthetic small molecules and combinatorial or naturally occurring compound libraries to discover novel therapeutics to regulate prostate function.

15

II. Specific Embodiments

A. The Proteins Associated with BPH

The present invention provides isolated proteins, allelic variants of the proteins, and conservative amino acid substitutions of the proteins. As used herein, the "protein" or "polypeptide" refers, in part, to a protein that has the human amino acid sequence depicted in SEQ ID NO: 2 or 4. The terms also refer to naturally occurring allelic variants and proteins that have a slightly different amino acid sequence than that specifically recited above. Allelic variants, though possessing a slightly different amino acid sequence than those recited above, will still have the same or similar biological functions associated with these proteins.

25

As used herein, the family of proteins related to the human amino acid sequence of SEQ ID NO: 2 or 4 refers to proteins that have been isolated from organisms in addition to humans. The methods used to identify and isolate other members of the family of proteins related to these proteins are described below.

30

The proteins of the present invention are preferably in isolated form. As used herein, a protein is said to be isolated when physical, mechanical or chemical methods are employed to remove the protein from cellular constituents that are normally associated with

the protein. A skilled artisan can readily employ standard purification methods to obtain an isolated protein.

The proteins of the present invention further include insertion, deletion or conservative amino acid substitution variants of the sequences set forth in SEQ ID NO: 2 or 4. As used herein, a conservative variant refers to alterations in the amino acid sequence that do not adversely affect the biological functions of the protein. A substitution, insertion or deletion is said to adversely affect the protein when the altered sequence prevents or disrupts a biological function associated with the protein. For example, the overall charge, structure or hydrophobic/hydrophilic properties of the protein can be altered without adversely affecting a biological activity. Accordingly, the amino acid sequence can be altered, for example to render the peptide more hydrophobic or hydrophilic, without adversely affecting the biological activities of the protein.

Ordinarily, the allelic variants, the conservative substitution variants, and the members of the protein family, will have an amino acid sequence having at least about 50%, 60%, 70% or 75% amino acid sequence identity with the sequences set forth in SEQ ID NO: 2 or 4, more preferably at least about 80%, even more preferably at least about 90%, and most preferably at least about 95%. Identity or homology with respect to such sequences is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the known peptides, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent homology, and not considering any conservative substitutions as part of the sequence identity (see section B for the relevant parameters). Fusion proteins, or N-terminal, C-terminal or internal extensions, deletions, or insertions into the peptide sequence shall not be construed as affecting homology.

Thus, the proteins of the present invention include molecules having the amino acid sequence disclosed in SEQ ID NO: 2 or 4; fragments thereof having a consecutive sequence of at least about 3, 4, 5, 6, 10, 15, 20, 25, 30, 35 or more amino acid residues of these proteins; amino acid sequence variants wherein one or more amino acid residues has been inserted N- or C-terminal to, or within, the disclosed coding sequence; and amino acid sequence variants of the disclosed sequence, or their fragments as defined above, that have been substituted by another residue. Such fragments, also referred to as peptides or polypeptides, may contain antigenic regions, functional regions of the protein identified as

regions of the amino acid sequence which correspond to known protein domains, as well as regions of pronounced hydrophilicity. The regions are all easily identifiable by using commonly available protein sequence analysis software such as MacVector® (Oxford Molecular).

5 Contemplated variants further include those containing predetermined mutations by, e.g., homologous recombination, site-directed or PCR mutagenesis, and the corresponding proteins of other animal species, including but not limited to rabbit, mouse, rat, porcine, bovine, ovine, equine and non-human primate species, and the alleles or other naturally occurring variants of the family of proteins; and derivatives wherein the protein has been
10 covalently modified by substitution, chemical, enzymatic, or other appropriate means with a moiety other than a naturally occurring amino acid (for example a detectable moiety such as an enzyme or radioisotope).

 The present invention further provides compositions comprising a protein or polypeptide of the invention and a diluent. Suitable diluents can be aqueous or non-aqueous
15 solvents or a combination thereof, and can comprise additional components, for example water-soluble salts or glycerol, that contribute to the stability, solubility, activity, and/or storage of the protein or polypeptide.

 As described below, members of the family of proteins can be used: (1) to identify agents which modulate the level of or at least one activity of the protein, (2) to identify
20 binding partners for the protein, (3) as an antigen to raise polyclonal or monoclonal antibodies, (4) as a therapeutic agent or target and (5) as a diagnostic agent or marker of BPH and other hyperplastic diseases.

B. Nucleic Acid Molecules

25 The present invention further provides nucleic acid molecules that encode the protein having SEQ ID NO: 2 or 4 and the related proteins herein described, preferably in isolated form. As used herein, "nucleic acid" is defined as RNA or DNA that encodes a protein or peptide as defined above, is complementary to a nucleic acid sequence encoding such peptides, hybridizes to the nucleic acid of SEQ ID NO: 1 or 3 and remains stably
30 bound to it under appropriate stringency conditions, encodes a polypeptide sharing at least about 35%, 40%, 50%, 60%, 70% or 75%, preferably at least about 80%, more preferably at least about 85%, and most preferably at least about 90% or 95% or more identity with the

peptide sequence of SEQ ID NO: 2 or 4 or exhibits at least about 35%, 40%, 50%, 60%, 70% or 75%, preferably at least about 80%, more preferably at least about 85%, and even more preferably at least about 90% or 95% or more nucleotide sequence identity over the open reading frames of SEQ ID NO: 1 or 3.

5 Specifically contemplated are genomic DNA, cDNA, mRNA and antisense molecules, as well as nucleic acids based on alternative backbones or including alternative bases, whether derived from natural sources or synthesized. Such hybridizing or complementary nucleic acids, however, are defined further as being novel and unobvious over any prior art nucleic acid including that which encodes, hybridizes under appropriate stringency conditions, or is complementary to nucleic acid encoding a protein according to
10 the present invention.

Homology or identity at the nucleotide or amino acid sequence level is determined by **BLAST** (**B**asic **L**ocal **A**lignment **S**earch **T**ool) analysis using the algorithm employed by the programs **blastp**, **blastn**, **blastx**, **tblastn** and **tblastx** (Altschul S.F. *et al.* (1997) *Nucleic*
15 *Acids Res* 25:3389-3402, and Karlin *et al.*, (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268, both fully incorporated by reference) which are tailored for sequence similarity searching. The approach used by the **BLAST** program is to first consider similar segments, with and without gaps, between a query sequence and a database sequence, then to evaluate the statistical significance of all matches that are identified and finally to summarize only
20 those matches which satisfy a preselected threshold of significance. For a discussion of basic issues in similarity searching of sequence databases, see Altschul *et al.*, (1994) *Nature Genetics* 6:119-129 which is fully incorporated by reference. The search parameters for **histogram**, **descriptions**, **alignments**, **expect** (*i.e.*, the statistical significance threshold for reporting matches against database sequences), **cutoff**, **matrix** and **filter** (low complexity)
25 are at the default settings. The default scoring matrix used by **blastp**, **blastx**, **tblastn**, and **tblastx** is the **BLOSUM62** matrix (Henikoff *et al.*, (1992) *Proc Natl Acad Sci USA* 89:10915-10919, fully incorporated by reference), recommended for query sequences over 85 nucleotides or amino acids in length.

For **blastn**, the scoring matrix is set by the ratios of **M** (*i.e.*, the reward score for a pair of matching residues) to **N** (*i.e.*, the penalty score for mismatching residues), wherein
30 the default values for **M** and **N** are 5 and -4, respectively. Four **blastn** parameters were adjusted as follows: **Q**=10 (gap creation penalty); **R**=10 (gap extension penalty); **wink**=1

(generates word hits at every $wink^{th}$ position along the query); and $gapw=16$ (sets the window width within which gapped alignments are generated). The equivalent **Blastp** parameter settings were $Q=9$; $R=2$; $wink=1$; and $gapw=32$. A **Bestfit** comparison between sequences, available in the GCG package version 10.0, uses DNA parameters $GAP=50$ (gap creation penalty) and $LEN=3$ (gap extension penalty) and the equivalent settings in protein comparisons are $GAP=8$ and $LEN=2$.

“Stringent conditions” are those that (1) employ low ionic strength and high temperature for washing, for example, 0.015 M NaCl/0.0015 M sodium citrate/0.1% SDS at 50°C, or (2) employ during hybridization a denaturing agent such as formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH 6.5 with 750 mM NaCl, 75 mM sodium citrate at 42°C. Another example is hybridization in 50% formamide, 5× SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5× Denhardt’s solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2× SSC and 0.1% SDS. A skilled artisan can readily determine and vary the stringency conditions appropriately to obtain a clear and detectable hybridization signal. Preferred molecules are those that hybridize under the above conditions to the complement of SEQ ID NO: 1 or 3 and which encode a functional protein. Even more preferred hybridizing molecules are those that hybridize under the above conditions to the complement strand of the open reading frame of SEQ ID NO: 1 or 3.

As used herein, a nucleic acid molecule is said to be “isolated” when the nucleic acid molecule is substantially separated from contaminant nucleic acid molecules encoding other polypeptides.

The present invention further provides fragments of the encoding nucleic acid molecule. As used herein, a fragment of an encoding nucleic acid molecule refers to a small portion of the entire protein coding sequence. The size of the fragment will be determined by the intended use. For example, if the fragment is chosen so as to encode an active portion of the protein, the fragment will need to be large enough to encode the functional region(s) of the protein. For instance, fragments which encode peptides corresponding to predicted antigenic regions may be prepared. If the fragment is to be used as a nucleic acid probe or PCR primer, then the fragment length is chosen so as to obtain a relatively small

number of false positives during probing/priming (see the discussion in Section H).

Fragments of the encoding nucleic acid molecules of the present invention (*i.e.*, synthetic oligonucleotides) that are used as probes or specific primers for the polymerase chain reaction (PCR), or to synthesize gene sequences encoding proteins of the invention, can easily be synthesized by chemical techniques, for example, the phosphotriester method of Matteucci *et al.*, (1981) (*J Am Chem Soc* 103:3185-3191) or using automated synthesis methods. In addition, larger DNA segments can readily be prepared by well known methods, such as synthesis of a group of oligonucleotides that define various modular segments of the gene, followed by ligation of oligonucleotides to build the complete modified gene.

The encoding nucleic acid molecules of the present invention may further be modified so as to contain a detectable label for diagnostic and probe purposes. A variety of such labels are known in the art and can readily be employed with the encoding molecules herein described. Suitable labels include, but are not limited to, biotin, radiolabeled nucleotides and the like. A skilled artisan can readily employ any such label to obtain labeled variants of the nucleic acid molecules of the invention.

Modifications to the primary structure itself by deletion, addition, or alteration of the amino acids incorporated into the protein sequence during translation can be made without destroying the activity of the protein. Such substitutions or other alterations result in proteins having an amino acid sequence encoded by a nucleic acid falling within the contemplated scope of the present invention.

C. Isolation of Other Related Nucleic Acid Molecules

As described above, the identification and characterization of the nucleic acid molecule having SEQ ID NO: 1 or 3 allows a skilled artisan to isolate nucleic acid molecules that encode other members of the protein family in addition to the sequences herein described.

For instance, a skilled artisan can readily use the amino acid sequence of SEQ ID NO: 2 or 4 to generate antibody probes to screen expression libraries prepared from appropriate cells. Typically, polyclonal antiserum from mammals such as rabbits immunized with the purified protein (as described below) or monoclonal antibodies can be used to probe a mammalian cDNA or genomic expression library, such as lambda gtl1

library, to obtain the appropriate coding sequence for other members of the protein family. The cloned cDNA sequence can be expressed as a fusion protein, expressed directly using its own control sequences, or expressed by constructions using control sequences appropriate to the particular host used for expression of the enzyme.

5 Alternatively, a portion of the coding sequence herein described can be synthesized and used as a probe to retrieve DNA encoding a member of the protein family from any mammalian organism. Oligomers containing approximately 18-20 nucleotides (encoding about a 6-7 amino acid stretch) are prepared and used to screen genomic DNA or cDNA libraries to obtain hybridization under stringent conditions or conditions of sufficient
10 stringency to eliminate an undue level of false positives.

Additionally, pairs of oligonucleotide primers can be prepared for use in a polymerase chain reaction (PCR) to selectively clone an encoding nucleic acid molecule. A PCR denature-anneal-extend cycle for using such PCR primers is well known in the art and can readily be adapted for use in isolating other encoding nucleic acid molecules.

15 Nucleic acid molecules encoding other members of the protein family may also be identified in existing genomic or other sequence information using any available computational method, including but not limited to: PSI-BLAST (Altschul, *et al.* (1997) *Nucleic Acids Res* 25:3389-3402); PHI-BLAST (Zhang, *et al.* (1998), *Nucleic Acids Res.* 26:3986-3990), 3D-PSSM (Kelly *et al.* (2000) *J Mol Biol* 299(2): 499-520); and other
20 computational analysis methods (Shi *et al.* (1999) *Biochem. Biophys. Res. Commun.* 262(1):132-8 and Matsunami *et. al.* (2000) *Nature* 404(6778):601-4.

D. rDNA molecules Containing a Nucleic Acid Molecule

The present invention further provides recombinant DNA molecules (rDNAs) that
25 contain a coding sequence. As used herein, a rDNA molecule is a DNA molecule that has been subjected to molecular manipulation *in situ*. Methods for generating rDNA molecules are well known in the art, for example, see Sambrook *et al.*, (1989) *Molecular Cloning - A Laboratory Manual*, Cold Spring Harbor Laboratory Press. In the preferred rDNA molecules, a coding DNA sequence is operably linked to expression control sequences
30 and/or vector sequences.

The choice of vector and/or expression control sequences to which one of the protein family encoding sequences of the present invention is operably linked depends -

directly, as is well known in the art, on the functional properties desired, *e.g.*, protein expression, and the host cell to be transformed. A vector contemplated by the present invention is at least capable of directing the replication or insertion into the host chromosome, and preferably also expression, of the structural gene included in the rDNA molecule.

Expression control elements that are used for regulating the expression of an operably linked protein encoding sequence are known in the art and include, but are not limited to, inducible promoters, constitutive promoters, secretion signals, and other regulatory elements. Preferably, the inducible promoter is readily controlled, such as being responsive to a nutrient in the host cell's medium.

In one embodiment, the vector containing a coding nucleic acid molecule will include a prokaryotic replicon, *i.e.*, a DNA sequence having the ability to direct autonomous replication and maintenance of the recombinant DNA molecule extrachromosomally in a prokaryotic host cell, such as a bacterial host cell, transformed therewith. Such replicons are well known in the art. In addition, vectors that include a prokaryotic replicon may also include a gene whose expression confers a detectable marker such as a drug resistance. Typical bacterial drug resistance genes are those that confer resistance to ampicillin or tetracycline.

Vectors that include a prokaryotic replicon can further include a prokaryotic or bacteriophage promoter capable of directing the expression (transcription and translation) of the coding gene sequences in a bacterial host cell, such as *E. coli*. A promoter is an expression control element formed by a DNA sequence that permits binding of RNA polymerase and transcription to occur. Promoter sequences compatible with bacterial hosts are typically provided in plasmid vectors containing convenient restriction sites for insertion of a DNA segment of the present invention. Typical of such vector plasmids are pUC8, pUC9, pBR322 and pBR329 available from BioRad Laboratories, (Richmond, CA), pPL and pKK223 available from Pharmacia (Piscataway, NJ).

Expression vectors compatible with eukaryotic cells, preferably those compatible with vertebrate cells, such as prostate cells, can also be used to form rDNA molecules that contain a coding sequence. Eukaryotic cell expression vectors, including viral vectors, are well known in the art and are available from several commercial sources. Typically, such vectors are provided containing convenient restriction sites for insertion of the desired DNA

segment. Typical of such vectors are pSVL and pKSV-10 (Pharmacia), pBPV-1/pML2d (International Biotechnologies, Inc.), pTDT1 (ATCC, #31255), the vector pCDM8 described herein, and the like eukaryotic expression vectors. Vectors may be modified to include prostate cell specific promoters if needed.

5 Eukaryotic cell expression vectors used to construct the rDNA molecules of the present invention may further include a selectable marker that is effective in an eukaryotic cell, preferably a drug resistance selection marker. A preferred drug resistance marker is the gene whose expression results in neomycin resistance, *i.e.*, the neomycin phosphotransferase (*neo*) gene. (Southern *et al.*, (1982) *J Mol Anal Genet* 1:327-341)

10 Alternatively, the selectable marker can be present on a separate plasmid, and the two vectors are introduced by co-transfection of the host cell, and selected by culturing in the appropriate drug for the selectable marker.

E. Host Cells Containing an Exogenously Supplied Coding Nucleic Acid Molecule

15 The present invention further provides host cells transformed with a nucleic acid molecule that encodes a protein of the present invention. The host cell can be either prokaryotic or eukaryotic. Eukaryotic cells useful for expression of a protein of the invention are not limited, so long as the cell line is compatible with cell culture methods and compatible with the propagation of the expression vector and expression of the gene
20 product. Preferred eukaryotic host cells include, but are not limited to, yeast, insect and mammalian cells, preferably vertebrate cells such as those from a mouse, rat, monkey or human cell line. Preferred eukaryotic host cells include Chinese hamster ovary (CHO) cells available from the ATCC as CCL61, NIH Swiss mouse embryo cells (NIH/3T3) available from the ATCC as CRL 1658, baby hamster kidney cells (BHK), and the like eukaryotic
25 tissue culture cell lines.

Any prokaryotic host can be used to express a rDNA molecule encoding a protein of the invention. The preferred prokaryotic host is *E. coli*.

Transformation of appropriate cell hosts with a rDNA molecule of the present invention is accomplished by well known methods that typically depend on the type of
30 vector used and host system employed. With regard to transformation of prokaryotic host cells, electroporation and salt treatment methods are typically employed, see, for example, Cohen *et al.*, (1972) *Proc Natl Acad Sci USA* 69:2110; and Sambrook *et al.*, (1989)

Molecular Cloning - A Laboratory Manual, Cold Spring Harbor Laboratory Press. With regard to transformation of vertebrate cells with vectors containing rDNAs, electroporation, cationic lipid or salt treatment methods are typically employed, see, for example, Graham *et al.*, (1973) *Virol* 52:456; Wigler *et al.*, (1979) *Proc Natl Acad Sci USA* 76:1373-1376.

5 Successfully transformed cells, *i.e.*, cells that contain a rDNA molecule of the present invention, can be identified by well known techniques including the selection for a selectable marker. For example, cells resulting from the introduction of an rDNA of the present invention can be cloned to produce single colonies. Cells from those colonies can be harvested, lysed and their DNA content examined for the presence of the rDNA using a
10 method such as that described by Southern (1975) *J. Mol. Biol.* 98:503 or Berent *et al.*, (1985) *Biotech* 3:208, or the proteins produced from the cell assayed via an immunological method.

F. Production of Recombinant Proteins using a rDNA Molecule

15 The present invention further provides methods for producing a protein of the invention using nucleic acid molecules herein described. In general terms, the production of a recombinant form of a protein typically involves the following steps:

 A nucleic acid molecule is first obtained that encodes a protein of the invention, such as a nucleic acid molecule comprising, consisting essentially of or consisting of SEQ
20 ID NO: 1 or 3, nucleotides 464-1450 of SEQ ID NO: 1, nucleotides 464-1453 of SEQ ID NO: 1, nucleotides 263-2188 of SEQ ID NO: 3 or nucleotides 263-2191 of SEQ ID NO: 3. If the encoding sequence is uninterrupted by introns, as are these open-reading-frames, it is directly suitable for expression in any host.

 The nucleic acid molecule is then preferably placed in operable linkage with suitable
25 control sequences, as described above, to form an expression unit containing the protein open reading frame. The expression unit is used to transform a suitable host and the transformed host is cultured under conditions that allow the production of the recombinant protein. Optionally the recombinant protein is isolated from the medium or from the cells; recovery and purification of the protein may not be necessary in some instances where some
30 impurities may be tolerated.

 Each of the foregoing steps can be done in a variety of ways. For example, the desired coding sequences may be obtained from genomic fragments and used directly in

appropriate hosts. The construction of expression vectors that are operable in a variety of hosts is accomplished using appropriate replicons and control sequences, as set forth above. The control sequences, expression vectors, and transformation methods are dependent on the type of host cell used to express the gene and were discussed in detail earlier. Suitable
5 restriction sites can, if not normally available, be added to the ends of the coding sequence so as to provide an excisable gene to insert into these vectors. A skilled artisan can readily adapt any host/expression system known in the art for use with the nucleic acid molecules of the invention to produce recombinant protein.

10 G. Methods to Identify Binding Partners

Another embodiment of the present invention provides methods for isolating and identifying binding partners of proteins of the invention. In general, a protein of the invention is mixed with a potential binding partner or an extract or fraction of a cell under conditions that allow the association of potential binding partners with the protein of the
15 invention. After mixing, peptides, polypeptides, proteins or other molecules that have become associated with a protein of the invention are separated from the mixture. The binding partner that bound to the protein of the invention can then be removed and further analyzed. To identify and isolate a binding partner, the entire protein, for instance a protein comprising the entire amino acid sequence of SEQ ID NO: 2 or 4 can be used.

20 Alternatively, a fragment of the protein can be used.

As used herein, a cellular extract refers to a preparation or fraction which is made from a lysed or disrupted cell. The preferred source of cellular extracts will be cells derived from human prostate tissue or cells, for instance, biopsy tissue or tissue culture cells from BPH. Alternatively, cellular extracts may be prepared from normal tissue or available cell
25 lines, particularly prostate derived cell lines.

A variety of methods can be used to obtain an extract of a cell. Cells can be disrupted using either physical or chemical disruption methods. Examples of physical disruption methods include, but are not limited to, sonication and mechanical shearing. Examples of chemical lysis methods include, but are not limited to, detergent lysis and
30 enzyme lysis. A skilled artisan can readily adapt methods for preparing cellular extracts in order to obtain extracts for use in the present methods.

Once an extract of a cell is prepared, the extract is mixed with the protein of the

invention under conditions in which association of the protein with the binding partner can occur. A variety of conditions can be used, the most preferred being conditions that closely resemble conditions found in the cytoplasm of a human cell. Features such as osmolarity, pH, temperature, and the concentration of cellular extract used, can be varied to optimize the association of the protein with the binding partner.

After mixing under appropriate conditions, the bound complex is separated from the mixture. A variety of techniques can be utilized to separate the mixture. For example, antibodies specific to a protein of the invention can be used to immunoprecipitate the binding partner complex. Alternatively, standard chemical separation techniques such as chromatography and density/sediment centrifugation can be used.

After removal of non-associated cellular constituents found in the extract, the binding partner can be dissociated from the complex using conventional methods. For example, dissociation can be accomplished by altering the salt concentration or pH of the mixture.

To aid in separating associated binding partner pairs from the mixed extract, the protein of the invention can be immobilized on a solid support. For example, the protein can be attached to a nitrocellulose matrix or acrylic beads. Attachment of the protein to a solid support aids in separating peptide/binding partner pairs from other constituents found in the extract. The identified binding partners can be either a single protein or a complex made up of two or more proteins. Alternatively, binding partners may be identified using a Far-Western assay according to the procedures of Takayama *et al.*, (1997) *Methods Mol Biol* 69:171-184 or Sauder *et al.*, (1996) *J. Gen Virol* 77:991-996 or identified through the use of epitope tagged proteins or GST fusion proteins.

Alternatively, the nucleic acid molecules of the invention can be used in a yeast two-hybrid system or other *in vivo* protein-protein detection system. The yeast two-hybrid system has been used to identify other protein partner pairs and can readily be adapted to employ the nucleic acid molecules herein described.

H. Methods to Identify Agents that Modulate the Expression a Nucleic Acid Encoding the Genes Associated with BPH

Another embodiment of the present invention provides methods for identifying agents that modulate the expression of a nucleic acid encoding a protein of the invention

such as a protein having the amino acid sequence of SEQ ID NO: 2 or 4. Such assays may utilize any available means of monitoring for changes in the expression level of the nucleic acids of the invention. As used herein, an agent is said to modulate the expression of a nucleic acid of the invention if it is capable of up- or down-regulating expression of the
5 nucleic acid in a cell.

In one assay format, cell lines that contain reporter gene fusions between nucleotides from within the open reading frame defined by nucleotides 464-1450 of SEQ ID NO: 1 or nucleotides 263-2188 of SEQ ID NO: 3 and/or the 5' and/or 3' regulatory elements and any assayable fusion partner may be prepared. Numerous assayable fusion partners are known
10 and readily available including the firefly luciferase gene and the gene encoding chloramphenicol acetyltransferase (Alam *et al.*, (1990) *Anal Biochem* 188:245-254). Cell lines containing the reporter gene fusions are then exposed to the agent to be tested under appropriate conditions and time. Differential expression of the reporter gene between samples exposed to the agent and control samples identifies agents which modulate the
15 expression of a nucleic acid of the invention.

Additional assay formats may be used to monitor the ability of the agent to modulate the expression of a nucleic acid encoding a protein of the invention, such as the protein having SEQ ID NO: 2 or 4. For instance, mRNA expression may be monitored directly by hybridization to the nucleic acids of the invention. Cell lines are exposed to the agent to be
20 tested under appropriate conditions and time and total RNA or mRNA is isolated by standard procedures such those disclosed in Sambrook *et al.*, (1989) *Molecular Cloning - A Laboratory Manual*, Cold Spring Harbor Laboratory Press).

The preferred cells will be those derived from human prostate tissue, for instance, prostate biopsy tissue or cultured prostate cells from normal or BPH patients, for example
25 BRF-55T cells (immortalized human prostate cells obtained from an individual with BPH; Iype, P.T. *et al.*, *Int J Oncol* (1998) 12:257-63 "Establishment and characterization of immortalized human cell lines from prostatic carcinoma and benign prostatic hyperplasia"). Alternatively, other available cells or cell lines may be used.

Probes to detect differences in RNA expression levels between cells exposed to the
30 agent and control cells may be prepared from the nucleic acids of the invention. It is preferable, but not necessary, to design probes which hybridize only with target nucleic acids under conditions of high stringency. Only highly complementary nucleic acid hybrids

form under conditions of high stringency. Accordingly, the stringency of the assay conditions determines the amount of complementarity which should exist between two nucleic acid strands in order to form a hybrid. Stringency should be chosen to maximize the difference in stability between the probe:target hybrid and probe:non-target hybrids.

5 Probes may be designed from the nucleic acids of the invention through methods known in the art. For instance, the G+C content of the probe and the probe length can affect probe binding to its target sequence. Methods to optimize probe specificity are commonly available in Sambrook *et al.*, (1989) *Molecular Cloning - A Laboratory Manual*, Cold Spring Harbor Laboratory Press) or Ausubel *et al.*, (1995) *Current Protocols in Molecular*
10 *Biology*, Greene Publishing Co.

Hybridization conditions are modified using known methods, such as those described by Sambrook *et al.* and Ausubel *et al.* as required for each probe. Hybridization of total cellular RNA or RNA enriched for polyA RNA can be accomplished in any available format. For instance, total cellular RNA or RNA enriched for polyA RNA can be
15 affixed to a solid support and the solid support exposed to at least one probe comprising at least one, or part of one of the sequences of the invention under conditions in which the probe will specifically hybridize. Alternatively, nucleic acid fragments comprising at least one, or part of one of the sequences of the invention can be affixed to a solid support, such as a silicon chip or a porous glass wafer. The solid support can then be exposed to total
20 cellular RNA or polyA RNA from a sample under conditions in which the affixed sequences will specifically hybridize. Such solid supports and hybridization methods are widely available, for example, those disclosed by Beattie, (1995) WO 95/11755. By examining for the ability of a given probe to specifically hybridize to an RNA sample from an untreated cell population and from a cell population exposed to the agent, agents which
25 up- or down-regulate the expression of a nucleic acid encoding the protein having the sequence of SEQ ID NO: 2 are identified.

Hybridization for qualitative and quantitative analysis of mRNAs may also be carried out by using a RNase Protection Assay (*i.e.*, RPA, see Ma *et al.* (1996) *Methods*
10:273-238). Briefly, an expression vehicle comprising cDNA encoding the gene product
30 and a phage specific DNA dependent RNA polymerase promoter (*e.g.*, T7, T3 or SP6 RNA polymerase) is linearized at the 3' end of the cDNA molecule, downstream from the phage promoter, wherein such a linearized molecule is subsequently used as a template for

synthesis of a labeled antisense transcript of the cDNA by *in vitro* transcription. The labeled transcript is then hybridized to a mixture of isolated RNA (*i.e.*, total or fractionated mRNA) by incubation at 45°C overnight in a buffer comprising 80 % formamide, 40 mM Pipes, pH 6.4, 0.4 M NaCl and 1 mM EDTA. The resulting hybrids are then digested in a buffer
5 comprising 40 µg/ml ribonuclease A and 2 µg/ml ribonuclease. After deactivation and extraction of extraneous proteins, the samples are loaded onto urea/polyacrylamide gels for analysis.

In another assay format, cells or cell lines are first identified which express the gene products of the invention physiologically. Cell and/or cell lines so identified would be
10 expected to comprise the necessary cellular machinery such that the fidelity of modulation of the transcriptional apparatus is maintained with regard to exogenous contact of agent with appropriate surface transduction mechanisms and/or the cytosolic cascades. Further, such cells or cell lines would be transduced or transfected with an expression vehicle (*e.g.*, a plasmid or viral vector) construct comprising an operable non-translated 5'-promoter
15 containing end of the structural gene encoding the instant gene products fused to one or more antigenic fragments, which are peculiar to the instant gene products, wherein said fragments are under the transcriptional control of said promoter and are expressed as polypeptides whose molecular weight can be distinguished from the naturally occurring polypeptides or may further comprise an immunologically distinct tag or other detectable
20 marker. Such a process is well known in the art (see Sambrook *et al.*, (1989)).

Cells or cell lines transduced or transfected as outlined above are then contacted with agents under appropriate conditions; for example, the agent in a pharmaceutically acceptable excipient is contacted with cells in an aqueous physiological buffer such as phosphate buffered saline (PBS) at physiological pH, Eagles balanced salt solution (BSS) at
25 physiological pH, PBS or BSS comprising serum or conditioned media comprising PBS or BSS and/or serum incubated at 37°C. Said conditions may be modulated as deemed necessary by one of skill in the art. Subsequent to contacting the cells with the agent, said cells will be disrupted and the polypeptides of the lysate are fractionated such that a polypeptide fraction is pooled and contacted with an antibody to be further processed by
30 immunological assay (*e.g.*, ELISA, immunoprecipitation or Western blot). The pool of proteins isolated from the "agent-contacted" sample will be compared with a control sample where only the excipient is contacted with the cells and an increase or decrease in the

immunologically generated signal from the "agent-contacted" sample compared to the control will be used to distinguish the effectiveness of the agent.

I. Methods to Identify Agents that Modulate the Level or at Least One Activity of the BPH Associated Proteins

Another embodiment of the present invention provides methods for identifying agents that modulate the level or at least one activity of a protein of the invention such as the protein having the amino acid sequence of SEQ ID NO: 2 or 4. Such methods or assays may utilize any means of monitoring or detecting the desired activity.

In one format, the relative amounts of a protein of the invention between a cell population that has been exposed to the agent to be tested compared to an un-exposed control cell population may be assayed. In this format, probes such as specific antibodies are used to monitor the differential expression of the protein in the different cell populations. Cell lines or populations are exposed to the agent to be tested under appropriate conditions and time. Cellular lysates may be prepared from the exposed cell line or population and a control, unexposed cell line or population. The cellular lysates are then analyzed with the probe.

Antibodies of the invention or antibody probes are prepared by immunizing suitable mammalian hosts in appropriate immunization protocols using the peptides, polypeptides or proteins of the invention if they are of sufficient length, or, if desired, or if required to enhance immunogenicity, conjugated to suitable carriers. Methods for preparing immunogenic conjugates with carriers such as BSA, KLH, or other carrier proteins are well known in the art. In some circumstances, direct conjugation using, for example, carbodiimide reagents may be effective; in other instances linking reagents such as those supplied by Pierce Chemical Co. (Rockford, IL), may be desirable to provide accessibility to the hapten. The hapten peptides can be extended at either the amino or carboxy terminus with a cysteine residue or interspersed with cysteine residues, for example, to facilitate linking to a carrier. Administration of the immunogens is conducted generally by injection over a suitable time period and with use of suitable adjuvants, as is generally understood in the art. During the immunization schedule, titers of antibodies are taken to determine adequacy of antibody formation.

While the polyclonal antisera produced in this way may be satisfactory for some applications, for pharmaceutical compositions, use of monoclonal preparations is preferred.

Immortalized cell lines which secrete the desired monoclonal antibodies may be prepared using the standard method of Kohler and Milstein (Nature (1975) 256:495-497) or modifications which effect immortalization of lymphocytes or spleen cells, as is generally known. The immortalized cell lines secreting the desired antibodies are screened by
5 immunoassay in which the antigen is the peptide hapten, polypeptide or protein. When the appropriate immortalized cell culture secreting the desired antibody is identified, the cells can be cultured either *in vitro* or by production in ascites fluid.

The desired monoclonal antibodies are then recovered from the culture supernatant or from the ascites supernatant. Fragments of the monoclonal antibodies or the polyclonal
10 antisera which contain the immunologically significant (antigen-binding) portion can be used as antagonists, as well as the intact antibodies. Use of immunologically reactive (antigen-binding) antibody fragments, such as the Fab, Fab', or F(ab')₂ fragments is often preferable, especially in a therapeutic context, as these fragments are generally less immunogenic than the whole immunoglobulin.

15 The antibodies or antigen-binding fragments may also be produced, using current technology, by recombinant means. Antibody regions that bind specifically to the desired regions of the protein can also be produced in the context of chimeras with multiple species origin, such as humanized antibodies.

Agents that are assayed in the above method can be randomly selected or rationally
20 selected or designed. As used herein, an agent is said to be randomly selected when the agent is chosen randomly without considering the specific sequences involved in the association of a protein of the invention alone or with its associated substrates, binding partners, etc. An example of randomly selected agents is the use a chemical library or a peptide combinatorial library, or a growth broth of an organism.

25 As used herein, an agent is said to be rationally selected or designed when the agent is chosen on a nonrandom basis which takes into account the sequence of the target site and/or its conformation in connection with the agent's action. Agents can be rationally selected or rationally designed by utilizing the peptide sequences that make up these sites. For example, a rationally selected peptide agent can be a peptide whose amino acid
30 sequence is identical to or a derivative of any functional consensus site.

The agents of the present invention can be, as examples, peptides, small molecules, vitamin derivatives, as well as carbohydrates. Dominant negative proteins, DNAs encoding

these proteins, antibodies to these proteins, peptide fragments of these proteins or mimics of these proteins may be introduced into cells to affect function. "Mimic" used herein refers to the modification of a region or several regions of a peptide molecule to provide a structure chemically different from the parent peptide but topographically and functionally similar to the parent peptide (see Grant GA. in: Meyers (ed.) Molecular Biology and Biotechnology (New York, VCH Publishers, 1995), pp. 659-664). A skilled artisan can readily recognize that there is no limit as to the structural nature of the agents of the present invention.

The peptide agents of the invention can be prepared using standard solid phase (or solution phase) peptide synthesis methods, as is known in the art. In addition, the DNA encoding these peptides may be synthesized using commercially available oligonucleotide synthesis instrumentation and produced recombinantly using standard recombinant production systems. The production using solid phase peptide synthesis is necessitated if non-gene-encoded amino acids are to be included.

Another class of agents of the present invention are antibodies immunoreactive with critical positions of proteins of the invention. Antibody agents are obtained by immunization of suitable mammalian subjects with peptides, containing as antigenic regions, those portions of the protein intended to be targeted by the antibodies.

J. Uses for Agents that Modulate the Expression or at least one Activity of the Proteins.

As provided in the Examples, the proteins and nucleic acids of the invention, such as the proteins having the amino acid sequence of SEQ ID NO: 2 or 4, are differentially expressed in BPH tissue. Agents that up- or down- regulate or modulate the expression of the protein or at least one activity of the protein, such as agonists or antagonists, of may be used to modulate biological and pathologic processes associated with the protein's function and activity.

As used herein, a subject can be any mammal, so long as the mammal is in need of modulation of a pathological or biological process mediated by a protein of the invention. The term "mammal" is defined as an individual belonging to the class Mammalia. The invention is particularly useful in the treatment of human subjects.

Pathological processes refer to a category of biological processes which produce a deleterious effect. For example, expression of a protein of the invention may be associated

with prostate cell growth or hyperplasia. As used herein, an agent is said to modulate a pathological process when the agent reduces the degree or severity of the process. For instance, BPH may be prevented or disease progression modulated by the administration of agents which up- or down-regulate or modulate in some way the expression or at least one activity of a protein of the invention.

The agents of the present invention can be provided alone, or in combination with other agents that modulate a particular pathological process. For example, an agent of the present invention can be administered in combination with other known drugs. As used herein, two agents are said to be administered in combination when the two agents are administered simultaneously or are administered independently in a fashion such that the agents will act at the same time.

The agents of the present invention can be administered via parenteral, subcutaneous, intravenous, intramuscular, intraperitoneal, transdermal, or buccal routes. Alternatively, or concurrently, administration may be by the oral route. The dosage administered will be dependent upon the age, health, and weight of the recipient, kind of concurrent treatment, if any, frequency of treatment, and the nature of the effect desired.

The present invention further provides compositions containing one or more agents which modulate expression or at least one activity of a protein of the invention. While individual needs vary, determination of optimal ranges of effective amounts of each component is within the skill of the art. Typical dosages comprise 0.1 to 100 $\mu\text{g/kg}$ body wt. The preferred dosages comprise 0.1 to 10 $\mu\text{g/kg}$ body wt. The most preferred dosages comprise 0.1 to 1 $\mu\text{g/kg}$ body wt.

In addition to the pharmacologically active agent, the compositions of the present invention may contain suitable pharmaceutically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically for delivery to the site of action. Suitable formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form, for example, water-soluble salts. In addition, suspensions of the active compounds as appropriate oily injection suspensions may be administered. Suitable lipophilic solvents or vehicles include fatty oils, for example, sesame oil, or synthetic fatty acid esters, for example, ethyl oleate or triglycerides. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension include, for example,

sodium carboxymethyl cellulose, sorbitol, and/or dextran. Optionally, the suspension may also contain stabilizers. Liposomes can also be used to encapsulate the agent for delivery into the cell.

5 The pharmaceutical formulation for systemic administration according to the invention may be formulated for enteral, parenteral or topical administration. Indeed, all three types of formulations may be used simultaneously to achieve systemic administration of the active ingredient.

10 Suitable formulations for oral administration include hard or soft gelatin capsules, pills, tablets, including coated tablets, elixirs, suspensions, syrups or inhalations and controlled release forms thereof.

15 In practicing the methods of this invention, the compounds of this invention may be used alone or in combination, or in combination with other therapeutic or diagnostic agents. In certain preferred embodiments, the compounds of this invention may be coadministered along with other compounds typically prescribed for these conditions according to generally accepted medical practice. The compounds of this invention can be utilized *in vivo*, ordinarily in mammals, such as humans, sheep, horses, cattle, pigs, dogs, cats, rats and mice, or *in vitro*.

K. Transgenic Animals

20 Transgenic animals containing mutant, knock-out or modified genes corresponding to the cDNA sequence of SEQ ID NO: 1 or 3, or the open reading frame encoding the polypeptide sequence of SEQ ID NO: 2 or 4 or fragments thereof having a consecutive sequence of at least about 3, 4, 5, 6, 10, 15, 20, 25, 30, 35 or more amino acid residues, are also included in the invention. Transgenic animals are genetically modified animals into
25 which recombinant, exogenous or cloned genetic material has been experimentally transferred. Such genetic material is often referred to as a "transgene." The nucleic acid sequence of the transgene, in this case a form of SEQ ID NO: 1 or 3, may be integrated either at a locus of a genome where that particular nucleic acid sequence is not otherwise normally found or at the normal locus for the transgene. The transgene may consist of
30 nucleic acid sequences derived from the genome of the same species or of a different species than the species of the target animal.

The term "germ cell line transgenic animal" refers to a transgenic animal in which

the genetic alteration or genetic information was introduced into a germ line cell, thereby conferring the ability of the transgenic animal to transfer the genetic information to offspring. If such offspring in fact possess some or all of that alteration or genetic information, then they too are transgenic animals.

5 The alteration or genetic information may be foreign to the species of animal to which the recipient belongs, foreign only to the particular individual recipient, or may be genetic information already possessed by the recipient. In the last case, the altered or introduced gene may be expressed differently than the native gene.

10 Transgenic animals can be produced by a variety of different methods including transfection, electroporation, microinjection, gene targeting in embryonic stem cells and recombinant viral and retroviral infection (*see, e.g.*, U.S. Patent No. 4,736,866; U.S. Patent No. 5,602,307; Mullins *et al.*, (1993) *Hypertension* 22:630-633; Brenin *et al.*, (1997) *Surg Oncol* 6:99-110; Tuan (1997) *Recombinant Gene Expression Protocols, Methods in Molecular Biology*, Humana Press).

15 A number of recombinant or transgenic mice have been produced, including those which express an activated oncogene sequence (U.S. Patent No. 4,736,866); express simian SV40 T-antigen (U.S. Patent No. 5,728,915); lack the expression of interferon regulatory factor 1 (IRF-1) (U.S. Patent No. 5,731,490); exhibit dopaminergic dysfunction (U.S. Patent No. 5,723,719); express at least one human gene which participates in blood pressure
20 control (U.S. Patent No. 5,731,489); display greater similarity to the conditions existing in naturally occurring Alzheimer's disease (U.S. Patent No. 5,720,936); have a reduced capacity to mediate cellular adhesion (U.S. Patent No. 5,602,307); possess a bovine growth hormone gene (Clutter *et al.*, (1996) *Genetics* 143:1753-1760); or, are capable of generating a fully human antibody response (McCarthy (1997) *Lancet* 349:405).

25 While mice and rats remain the animals of choice for most transgenic experimentation, in some instances it is preferable or even necessary to use alternative animal species. Transgenic procedures have been successfully utilized in a variety of non-murine animals, including sheep, goats, pigs, dogs, cats, monkeys, chimpanzees, hamsters, rabbits, cows and guinea pigs (*see, e.g.*, Kim *et al.*, (1997) *Mol Reprod Dev* 46:515-526;
30 Houdebine (1995) *Reprod Nutr Dev* 35:609-617; Petters (1994) *Reprod Fertil Dev* 6:643-645; Schnieke *et al.*, (1997) *Science* 278:2130-2133; and Amoah (1997) *J Animal Science* 75:578-585).

The method of introduction of nucleic acid fragments into recombination competent mammalian cells can be by any method which favors co-transformation of multiple nucleic acid molecules. Detailed procedures for producing transgenic animals are readily available to one skilled in the art, including the disclosures in U.S. Patent No. 5,489,743 and U.S.

5 Patent No. 5,602,307.

L. Diagnostic Methods

As the genes and proteins of the invention are differentially expressed in BPH tissue compared to normal prostate tissue, the genes and proteins of the invention may be used to
10 diagnose or monitor BPH, prostate function, or to track disease progression. One means of diagnosing BPH using the nucleic acid molecules or proteins of the invention involves obtaining prostate tissue from living subjects. Obtaining tissue samples from living sources is problematic for tissues such as prostate. However, due to the nature of the treatment paradigms for BPH, biopsy may be necessary. When possible, urine, blood or peripheral
15 lymphocyte samples may be used as the tissue sample in the assay. Commonly, in hyperplastic diseases, genes which are up-regulated in the affected tissue (for example the prostate) are also up-regulated in lymphocytes, which may be isolated from whole blood.

The use of molecular biological tools has become routine in forensic technology. For example, nucleic acid probes comprising all or at least part of the sequence of SEQ ID
20 NO: 1 may be used to determine the expression of a nucleic acid molecule in forensic/pathology specimens. Further, nucleic acid assays may be carried out by any means of conducting a transcriptional profiling analysis. In addition to nucleic acid analysis, forensic methods of the invention may target the proteins of the invention, particularly a protein comprising SEQ ID NO: 2 or 4, to determine up- or down-regulation
25 of the genes (Shiverick *et al.*, (1975) *Biochim Biophys Acta* 393:124-133).

Methods of the invention may involve treatment of tissues with collagenases or other proteases to make the tissue amenable to cell lysis (Semenov DE *et al.*, (1987) *Biull Eksp Biol Med* 104:113-116). Further, it is possible to obtain biopsy samples from different regions of the prostate for analysis.

30 Assays to detect nucleic acid or protein molecules of the invention may be in any available format. Typical assays for nucleic acid molecules include hybridization or PCR based formats. Typical assays for the detection of proteins, polypeptides or peptides of the

invention include the use of antibody probes in any available format such as *in situ* binding assays, etc. See Harlow & Lane, (1988) Antibodies - A Laboratory Manual, Cold Spring Harbor Laboratory Press. In preferred embodiments, assays are carried-out with appropriate controls.

5 The above methods may also be used in other diagnostic protocols, including protocols and methods to detect disease states in other tissues or organs, for example in tissues in which the gene is detected.

 Without further description, it is believed that one of ordinary skill in the art can, using the preceding description and the following illustrative examples, make and utilize the
10 compounds of the present invention and practice the claimed methods. The following working examples therefore, specifically point out preferred embodiments of the present invention, and are not to be construed as limiting in any way the remainder of the disclosure.

15 **EXAMPLES**

Example 1

Identification of Differentially Expressed BPH mRNA

 Human tissue was obtained from the transitional zone of the prostate in biopsy samples from normal individuals and from patients with BPH or prostate cancer. BPH was
20 defined histologically in all samples. Normal tissue and asymptomatic BPH samples came from individuals who died of trauma, and did not report symptoms. Patients having BPH with symptoms were defined as those with a need for frequent urination; in these patients, a radical prostatectomy had been performed. Prostate cancer patients provided age-matched tissue samples for symptomatic BPH patients, but were without symptoms and without
25 cancer in the transitional zone under histological examination.

 Microarray sample preparation was conducted with minor modifications, following the protocols set forth in the Affymetrix GeneChip® Expression Analysis Manual. Frozen tissue was ground to a powder using a Spex Certiprep 6800 Freezer Mill. Total RNA was extracted with Trizol® (GibcoBRL) utilizing the manufacturer's protocol. The total RNA
30 yield for each sample was 200-500 µg per 300 mg tissue weight. mRNA was isolated using the Oligotex mRNA Midi kit® (Qiagen) followed by ethanol precipitation. Double stranded cDNA was generated from mRNA using the SuperScript Choice® system (GibcoBRL).

First strand cDNA synthesis was primed with a T7-(dT₂₄) oligonucleotide. The cDNA was phenol-chloroform extracted and ethanol precipitated to a final concentration of 1 mg/ml. From 2 mg of cDNA, cRNA was synthesized using Ambion's T7 MegaScript *in vitro* Transcription Kit.

- 5 To biotin label the cRNA, nucleotides Bio-11-CTP and Bio-16-UTP (Enzo Diagnostics) were added into the reaction. Following a 37°C incubation for six hours, impurities were removed from the labeled cRNA following the Rneasy Mini kit protocol (Qiagen). cRNA was fragmented (5' fragmentation buffer consisting of 200 mM Tris-acetate (pH 8.1), 500 mM KOAc, 150 mM MgOAc) for thirty-five minutes at 94°C.
- 10 Following the Affymetrix protocol, 55 µg of fragmented cRNA was hybridized on a Human chip set and the HuGeneFL array for twenty-four hours at 60 rpm in a 45°C hybridization oven. The chips were washed and stained with Streptavidin Phycoerythrin (SAPE) (Molecular Probes) in Affymetrix fluidics stations. To amplify staining, SAPE solution was added twice with an anti-streptavidin biotinylated antibody (Vector Laboratories) staining
- 15 step in between. Hybridization to the probe arrays was detected by fluorometric scanning (Hewlett Packard Gene Array Scanner). Data was analyzed using Affymetrix GeneChip (version 3.0) and Expression Data Mining Tool (EDMT) software (version 1.0).

- Differential expression of genes between the BPH and normal prostate samples was determined using the Affymetrix GeneChip 42K human gene chip set by the following
- 20 criteria: (1) For each gene, Affymetrix GeneChip average difference values were determined by standard Affymetrix EDT software algorithms, which also made "Absent" (=not detected), "Present" (=detected) or "Marginal" (=not clearly Absent or Present) calls for each GeneChip element; (2) all negative values (=Absent) were raised to a floor of +20 (positive 20) so that fold change calculations could be made where values were not already
- 25 greater than or equal to +20; (3) median levels of expression were compared between the normal control group and the BPH with symptoms disease group to obtain greater than or equal to 3-fold up/down values; (4) The median value for the higher expressing group needed to be greater than or equal to 200 average difference units in order to be considered for statistical significance; (5) Genes passing the criteria of #1-4 were analyzed for
- 30 statistical significance using a two-tailed T test and deemed statistically significant if $p < 0.05$.

Figure 1 shows the expression of JT460914 across 25 human samples, including 10 normals, 5 BPH samples from individuals without symptoms, 5 BPH samples from individuals who were prostate cancer patients, and 5 BPH samples from patients with symptoms. A significant decrease was observed (-15.8-fold; $p=0.01366$) in average JT460914 expression levels in tissue from BPH patients with symptoms compared to normal prostate tissue samples (Normal 1 and Normal 2 groups), indicating that down-regulation of JT460914 may be diagnostic for BPH in symptomatic patients. Down-regulation of expression is observed in samples from people with BPH and cancer as well as from people with BPH and symptoms. The expression data show that down-regulation of JT460914 is diagnostic for BPH in general and more strongly for BPH in cancer patients and BPH in patients with symptoms.

Figure 4 shows the expression of JT156897 across 25 human samples, including 10 normals, 5 BPH samples from individuals without symptoms, 5 BPH samples from individuals who were prostate cancer patients, and 5 BPH samples from patients with symptoms. In this case, however, a significant increase was observed (2.4-fold; $p=0.033975$) in average JT156897 expression levels in tissue from BPH patients with symptoms compared to normal prostate tissue samples (Normal 1 and Normal 2 groups), indicating that up-regulation of JT156897 may be diagnostic for BPH in symptomatic patients.

Example 2

Cloning of Full Length Human cDNAs Corresponding to the differentially expressed mRNA species and Comparative Expression Studies

The full length cDNA having SEQ ID NO: 1 or 3 was obtained by the oligo-pulling method. Briefly, a gene-specific oligo was designed based on the sequence of JT460914 or JT156897. The oligo was labeled with biotin and used to hybridize with 2 μ g of single strand plasmid DNA (cDNA recombinants) from a placental cDNA library following the procedures of Sambrook *et al.* The hybridized cDNAs were separated by streptavidin-conjugated beads and eluted by heating. The eluted cDNA was converted to double strand plasmid DNA and used to transform *E. coli* cells (DH10B) and the longest cDNA was screened. After positive selection was confirmed by PCR using gene-specific primers, the cDNA clone was subjected to DNA sequencing.

The nucleotide sequence of the full-length human cDNAs corresponding to the differentially regulated mRNAs detected above is set forth in SEQ ID NOS: 1 and 3. The cDNA of SEQ ID NO: 1 comprises 3398 base pairs (3376 base pairs and a polyA tail). An open reading frame within the cDNA nucleotide sequence of SEQ ID NO: 1, at nucleotides 464-1450 (464-1453 including the stop codon), encodes a protein of 329 amino acids. The amino acid sequence corresponding to a predicted protein encoded by SEQ ID NO: 1 is set forth in SEQ ID NO: 2. The cDNA of SEQ ID NO: 3 comprises 3357 base pairs (3337 base pairs and a polyA tail). An open reading frame within the cDNA nucleotide sequence of SEQ ID NO: 1, at nucleotides 263-2188 (263-2191 including the stop codon), encodes a protein of 642 amino acids. The amino acid sequence corresponding to a predicted protein encoded by SEQ ID NO: 3 is set forth in SEQ ID NO: 4.

The lower part of Figure 1 shows an electronic Northern, in which the expression level of JT460914 was measured across a panel of normal tissues using the Affymetrix 42K human GeneChip set. For each tissue type, the mean \pm SDM is shown as a horizontal bar graph for samples obtained from 3 or more normal individuals. The highest expression level in normal tissue was observed in muscle and prostate, with lower but significant expression also observed in stomach. Thus, based upon the tissue panel tested, expression in normal tissue is quite strongly restricted. Because BPH with symptoms is associated with hypertrophic growth of the prostate and/or a low grade chronic inflammatory response, down-regulation of JT460914 DNA, or of protein products derived from JT460914, is likely to be of diagnostic value in disease states involving inflammatory responses or hypertrophic growth in tissues in which JT460914 is found (prostate, muscle or stomach). JT460914 nucleotides or protein products derived from JT460914 are potential targets for therapeutic intervention.

The lower part of Figure 4 shows an electronic Northern, in which the expression level of JT156897 was measured across a panel of normal tissues using the Affymetrix 42K human GeneChip set. For each tissue type, the mean \pm SDM is shown as a horizontal bar graph for samples obtained from three or more normal individuals. The highest expression level in normal tissue was observed in breast and cervix, but relatively strong expression was also observed in muscle, myometrium, endometrium, stomach, small intestine, lung, colon, and esophagus. In fact, expression was observed in essentially all of the tissues shown in the lower half of Figure 4 with brain the lowest. Because BPH with symptoms is

associated with hypertrophic growth of the prostate and/or a low grade chronic inflammation, upregulated expression of JT156897 in disease states involving hypertrophic growth and/or inflammation in general is likely of diagnostic value. This would apply to all of the tissues shown in the electronic Northern in lower half of Figure 4.

5 Figures 2 and 5 show the results of the quantitative PCR analysis of expression levels of mRNA corresponding to SEQ ID NO: 1 and 3 in various human tissue samples. Real time PCR detection was accomplished by the use of the ABI PRISM 7700 Sequence Detection System. The 7700 measures the fluorescence intensity of the sample each cycle and is able to detect the presence of specific amplicons within the PCR reaction. Each
10 sample was assayed for the level of GAPDH and mRNA corresponding to SEQ ID NO: 1 or 3. GAPDH detection was performed using Perkin Elmer part #402869 according to the manufacturer's directions. Primers were designed from SEQ ID NO: 1 or 3 using Primer Express, a program developed by PE to efficiently find primers and probes for specific sequences. These primers were used in conjunction with SYBR green (Molecular Probes),
15 a nonspecific double stranded DNA dye, to measure the expression level mRNA corresponding to SEQ ID NO: 1 or 3, which was normalized to the GAPDH level in each sample.

 Figures 3 and 6 show the results of hydrophobicity analyses of the amino acid sequences of SEQ ID NO: 2 and 4, respectively.

20

Example 3

Detection of mRNA for BPH Disease Screening

 The expression level of mRNA corresponding to SEQ ID NO: 1 or 3 is determined in prostate tissue biopsy samples, in urine or in lymphocytes from blood samples, as
25 described in Example 1 and Figure 1, *i.e.*, by screening mRNA samples on a GeneChip, or as described in Example 2 and Figure 2, *i.e.*, by semi-quantitative PCR analysis using the fluorescent detection system. Alternatively, samples from non-prostate hyperplastic tissues in malignant or non-malignant states may also be analyzed. Tissue samples from patients with BPH with symptoms and from normal subjects may be used as positive and negative
30 controls. Using standard assay methods, a level of expression of JT460914 lower than that of the normal control, or a level of expression of JT156897 higher than that of the normal control, is indicative of BPH or a likelihood of developing BPH.

Alternatively, because BPH with symptoms involves a chronic inflammatory response, the down-regulation of JT460914 or the up-regulation of JT156897 in prostate tissue could specifically identify cases in which inflammation is occurring, which would have additional diagnostic and prognostic implications and utility. More generally for
5 inflammatory conditions, the tissue distribution results (see Figures 1 and 4) indicate that the down-regulation of JT460914 or the up-regulation of JT156897 is likely to be of diagnostic significance in inflammatory responses in tissues in which these genes are found and a potential targets for therapeutic intervention.

Although the present invention has been described in detail with reference to
10 examples above, it is understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the following claims. All cited patents, patent applications and publications referred to in this application are herein incorporated by reference in their entirety.

-34-

What is claimed:

1. An isolated nucleic acid molecule selected from the group consisting of: (a) an isolated nucleic acid molecule that encodes the amino acid sequence of SEQ ID NO: 2 or 4;
5 (b) an isolated nucleic acid molecule which hybridizes to a nucleic acid molecule comprising SEQ ID NO: 1 or 3 and which encodes the amino acid sequence of SEQ ID NO: 2 or 4 or a variant thereof; and (c) an isolated nucleic acid molecule comprising at least about 60% sequence identity to the entire open reading frame of SEQ ID NO: 1 or 3..

10 2. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule comprises nucleotides 464-1450 of SEQ ID NO: 1.

15 3. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule consists of nucleotides 464-1450 of SEQ ID NO: 1.

4. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule comprises nucleotides 464-1453 of SEQ ID NO: 1.

20 5. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule comprises nucleotides 263-2188 of SEQ ID NO: 3.

6. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule consists of nucleotides 263-2188 of SEQ ID NO: 3.

25 7. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule comprises nucleotides 263-2191 of SEQ ID NO: 3.

8. The isolated nucleic acid molecule of any one of claims 1-7, wherein said nucleic acid molecule is operably linked to one or more expression control elements.

30

9. A vector comprising an isolated nucleic acid molecule of any one of claims 1-7.

10. A host cell transformed to contain the nucleic acid molecule of any one of claims 1-7.

11. A host cell comprising a vector of claim 9.

12. A host cell of claim 11, wherein said host is selected from the group consisting of prokaryotic hosts and eukaryotic hosts.

13. A method for producing a polypeptide comprising the step of culturing a host cell transformed with the nucleic acid molecule of any one of claims 1-7 under conditions in which the protein encoded by said nucleic acid molecule is expressed.

14. The method of claim 13, wherein said host cell is selected from the group consisting of prokaryotic hosts and eukaryotic hosts.

15. An isolated polypeptide produced by the method of claim 13.

16. An isolated polypeptide or protein selected from the group consisting of an isolated polypeptide comprising the amino acid sequence of SEQ ID NO: 2 or 4, an isolated polypeptide comprising conservative amino acid substitutions of SEQ ID NO: 2 or 4, an isolated polypeptide comprising naturally occurring amino acid sequence variants of SEQ ID NO: 2 or 4, and an isolated polypeptide exhibiting at least about 60% amino acid sequence identity over the entire length of SEQ ID NO: 2 or 4.

17. An isolated antibody or antigen-binding antibody fragment that binds to a polypeptide of either claim 15 or 16.

18. The antibody or antibody fragment of claim 17 wherein said antibody or antibody fragment is a monoclonal or polyclonal antibody.

19. A method of identifying an agent which modulates the expression of a nucleic acid encoding a protein of claim 16, comprising the steps of:

(a) exposing cells which express the nucleic acid to the agent; and
(b) determining whether the agent modulates expression of said nucleic acid,
thereby identifying an agent which modulates the expression of a nucleic acid encoding the
protein.

5

20. A method of identifying an agent which modulates at least one activity of a
protein of claim 16, comprising the steps of:

(a) exposing cells which express the protein to the agent;
(b) determining whether the agent modulates at least one activity of said protein,
10 thereby identifying an agent which modulates at least one activity of the protein.

21. The method of claim 20, wherein the agent modulates one activity of the
protein.

15 22. A method of identifying binding partners for a protein of claim 16, comprising
the steps of:

(a) exposing said protein to a potential binding partner; and
(b) determining if the potential binding partner binds to said protein, thereby
identifying binding partners for the protein.

20

23. A method of modulating the expression of a nucleic acid encoding a protein of
claim 16, comprising the step of:

(a) administering an effective amount of an agent which modulates the expression
of a nucleic acid encoding the protein.

25

24. A method of modulating at least one activity of a protein of claim 16,
comprising the step of:

(a) administering an effective amount of an agent which modulates at least one
activity of the protein.

30

25. A non-human transgenic animal modified to contain a nucleic acid molecule of
any of claims 1-7.

26. The transgenic animal of claim 25, wherein the nucleic acid molecule contains a mutation that prevents expression of the encoded protein.

5 27. A method of diagnosing a disease state in a subject, comprising the step of determining the level of expression of a nucleic acid molecule or protein of any one of claims 1-7 or 16.

10 28. The method of claim 27, wherein the disease state is benign prostatic hyperplasia.

 29. The method of claim 28, wherein the disease state is benign prostatic hyperplasia with symptoms.

15 30. A composition comprising a diluent and a polypeptide or protein selected from the group consisting of an isolated polypeptide comprising the amino acid sequence of SEQ ID NO: 2 or 4; an isolated polypeptide comprising a fragment of at least 6 amino acids of SEQ ID NO: 2 or 4, an isolated polypeptide comprising conservative amino acid
20 substitutions of SEQ ID NO: 2 or 4, an isolated polypeptide comprising naturally occurring amino acid sequence variants of SEQ ID NO: 2 or 4 and an isolated polypeptide exhibiting at least about 60% amino acid sequence identity with SEQ ID NOS: 2 or 4.

FIG. 1a



FIG. 1b

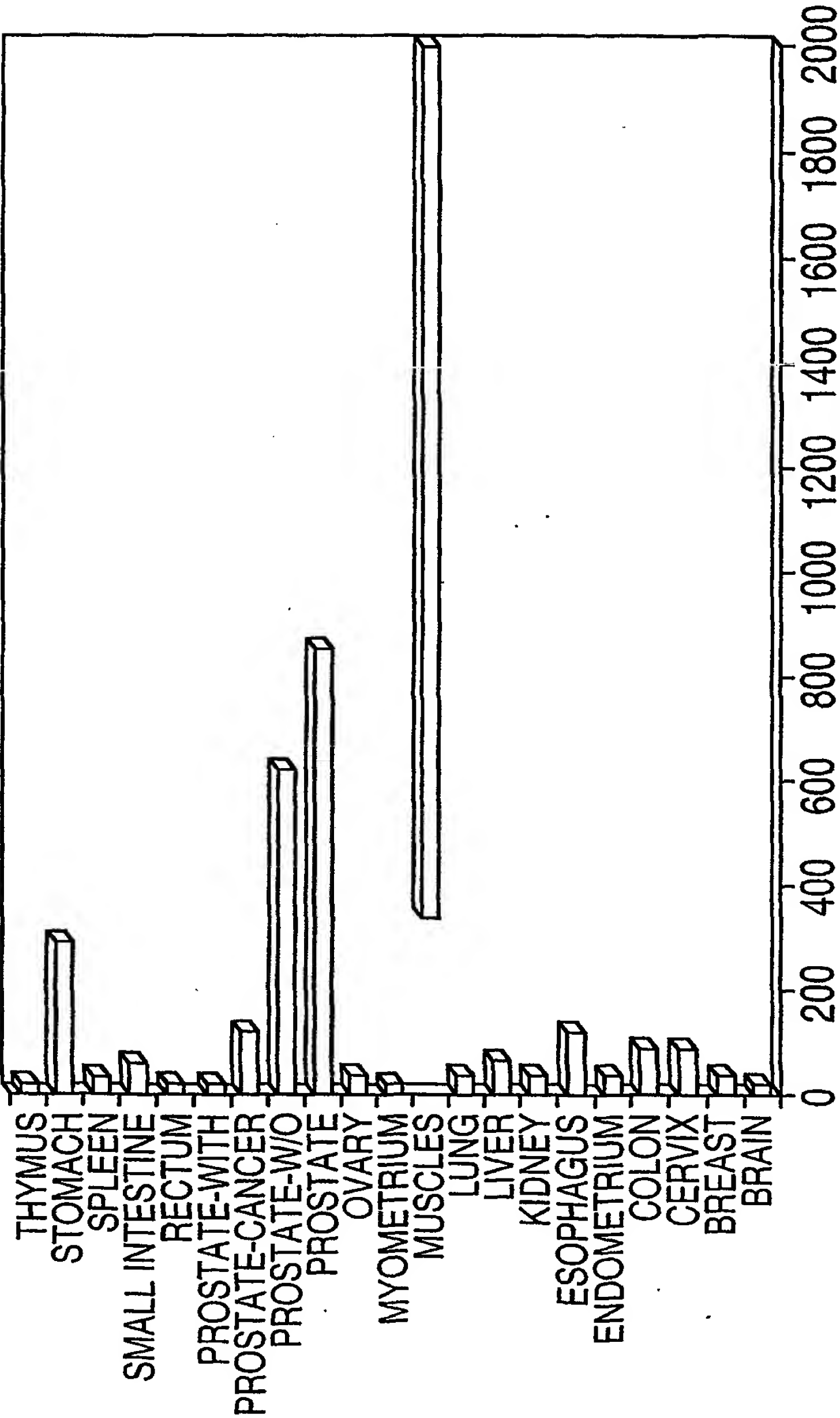


FIG. 2

AA460914

25

30

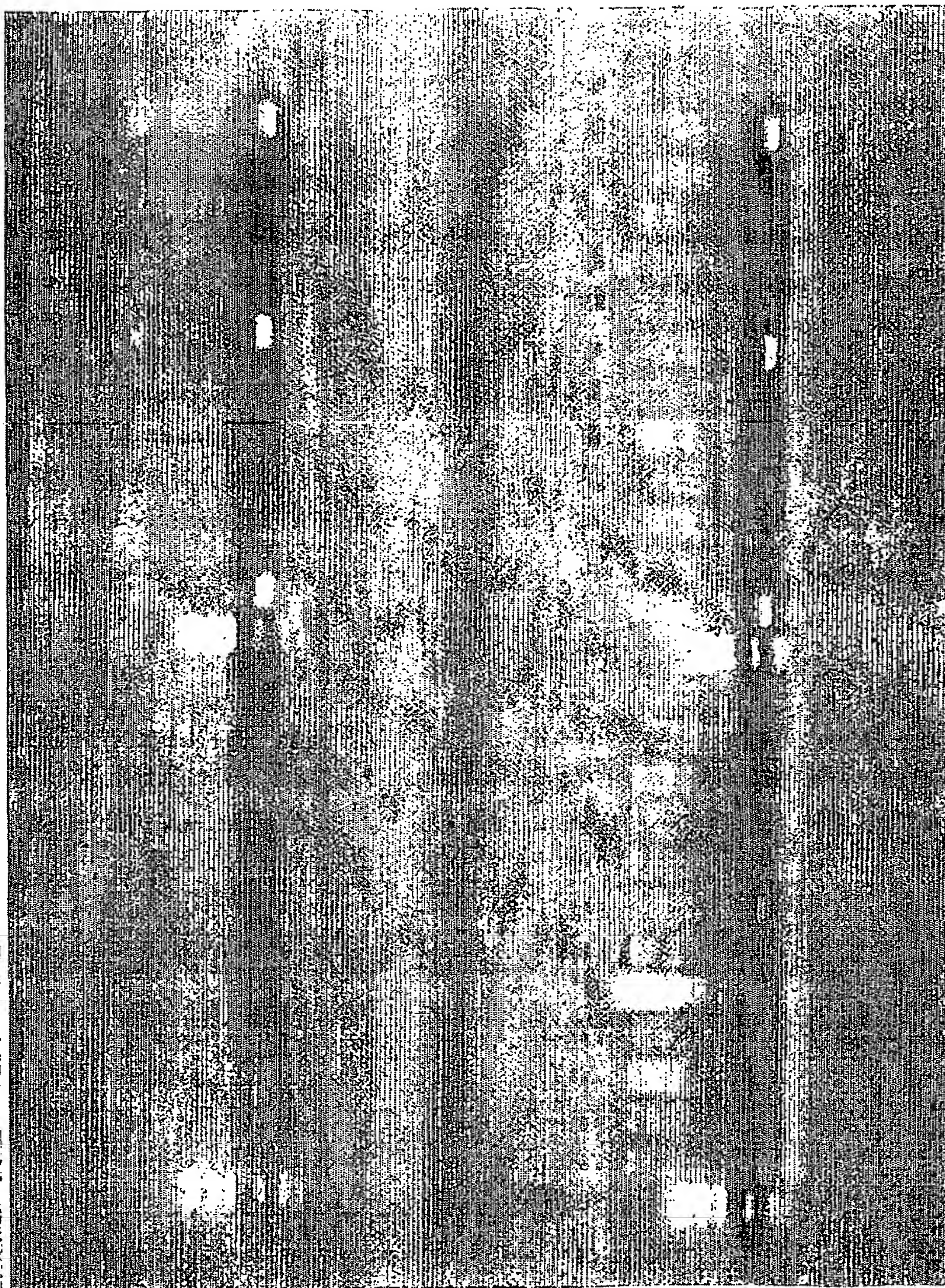


FIG. 3

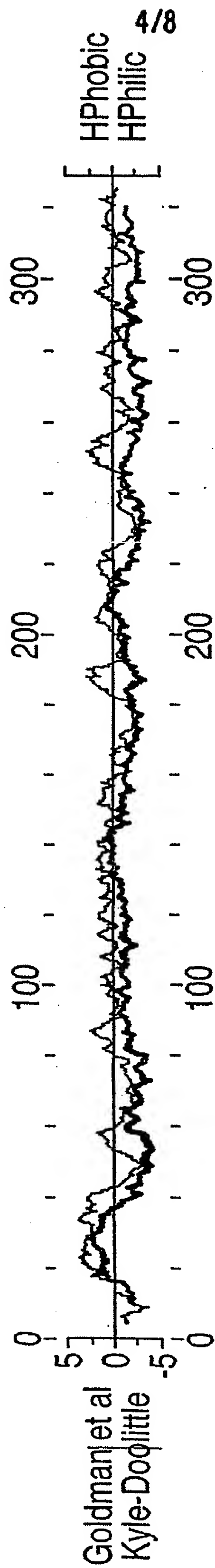


FIG. 4a

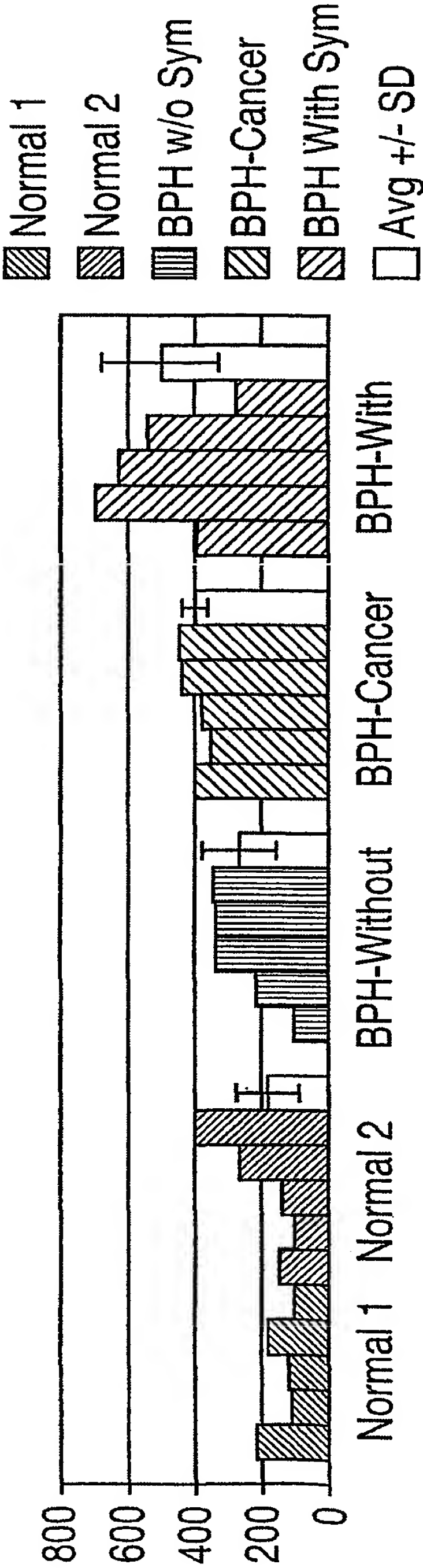


FIG. 5

AA156897

7/8

25

30

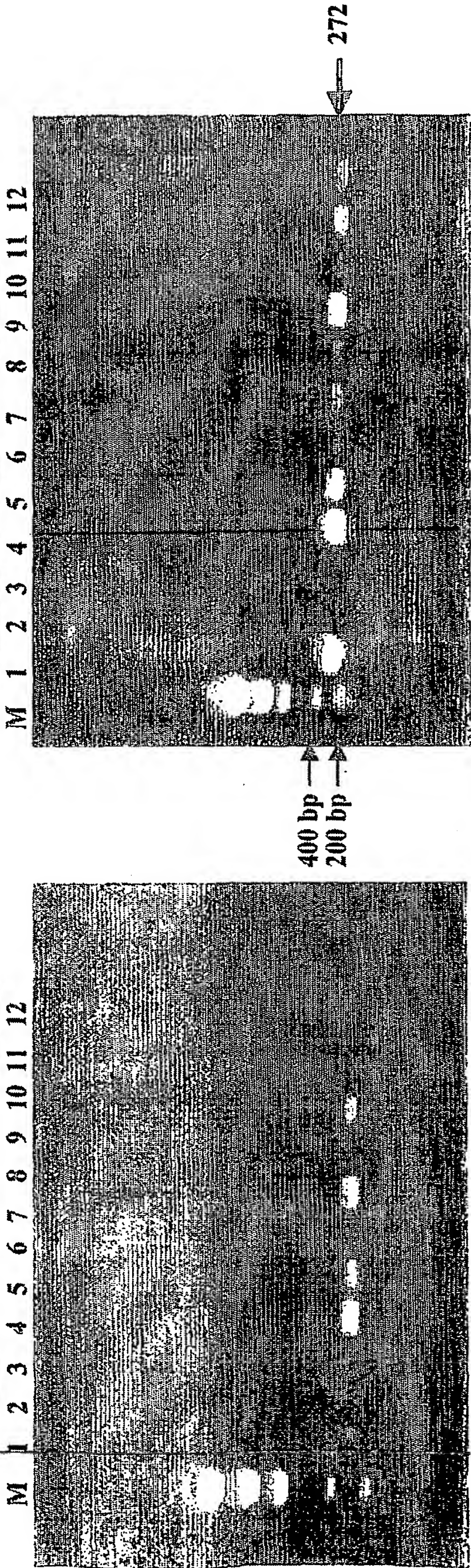
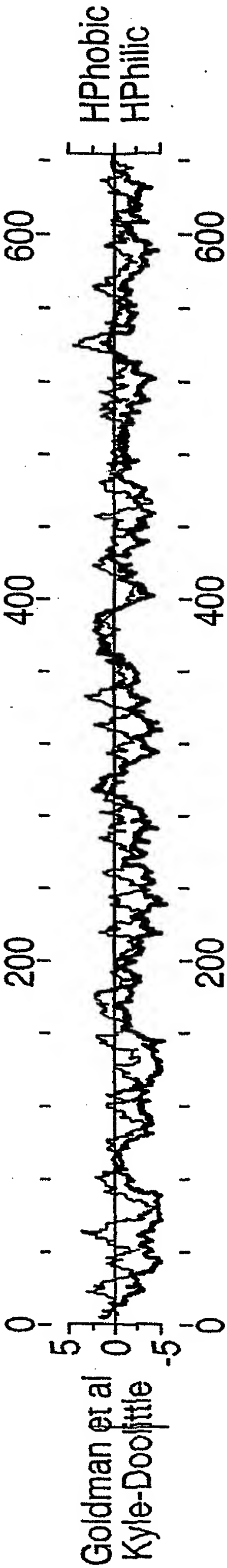


FIG. 6



SEQUENCE LISTING

<110> Munger, William E.
Kulkarni, Prakash
Getzenberg, Robert H.

<120> Identification of cDNAs Associated with Benign
Prostatic Hyperplasia

<130> 44921-5030-WO

<140>

<141>

<150> US 60/212,726

<151> 2000-06-20

<160> 4

<170> PatentIn Ver. 2.1

<210> 1

<211> 3398

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (464)..(1453)

<223> Clone No. JT460914

<400> 1

```

gcaaaacaaa gtgcaggatc gagagaaaat cgattgcttg aaaaactggg tatttttactt 60
gctgcaagtc ccgactgtgt cagctgatga gtttaacact atgacagcac agtcctcacc 120
tgcctgctcg ctgctgtaat cccacactg ttgtgtgcc tatagcacct ggcacacggg 180
aagcgctcaa gagatgtgtg tgaatgaagg agccacatct tgacttgcca cgcggagcag 240
taagctgcag ttatttggtg gcggtggaag aataaattct acctcctgct tgcctcgctg 300
cagtggggat gagggagagg ggactcagcc agatgctgaa ttccccagac agctgacagg 360
ttcccaattt ccaggaccag ctgttcagga gcatcccgga cgagagccag gcacatttga 420
gacttggatc caactaaaga ccgcccagga ttcttctgca gca atg tcg gtg tta 475
                                     Met Ser Val Leu
                                     1

gaa gaa aat cgg ccg ttt gct caa caa tta tcc aat gtc tac ttt aca 523
Glu Glu Asn Arg Pro Phe Ala Gln Gln Leu Ser Asn Val Tyr Phe Thr
  5              10              15              20

ata ctt tcg ctg ttc tgt ttt aag ctt ttt gtg aaa atc agc ctt gcc 571
Ile Leu Ser Leu Phe Cys Phe Lys Leu Phe Val Lys Ile Ser Leu Ala

```


25										30					35					
atc	ctc	agt	cat	ttc	tac	ata	gtg	aaa	ggc	aac	cgc	aag	gaa	gcg	gca	619				
Ile	Leu	Ser	His	Phe	Tyr	Ile	Val	Lys	Gly	Asn	Arg	Lys	Glu	Ala	Ala					
			40					45					50							
agg	ata	gca	gct	gaa	ttt	tat	gga	gta	acc	caa	gga	caa	ggg	tcc	tgg	667				
Arg	Ile	Ala	Ala	Glu	Phe	Tyr	Gly	Val	Thr	Gln	Gly	Gln	Gly	Ser	Trp					
		55					60					65								
gca	gat	cga	tca	cca	cta	cat	gaa	gca	gca	agt	caa	ggg	cgc	ctt	ctt	715				
Ala	Asp	Arg	Ser	Pro	Leu	His	Glu	Ala	Ala	Ser	Gln	Gly	Arg	Leu	Leu					
	70					75					80									
gct	ctg	aga	aca	tta	tta	tca	cag	ggg	tat	aat	gta	aat	gca	gta	acc	763				
Ala	Leu	Arg	Thr	Leu	Leu	Ser	Gln	Gly	Tyr	Asn	Val	Asn	Ala	Val	Thr					
85					90					95					100					
tta	gac	cat	gtc	acc	cca	ttg	cac	gaa	gcc	tgc	ctt	gga	gat	cac	gtg	811				
Leu	Asp	His	Val	Thr	Pro	Leu	His	Glu	Ala	Cys	Leu	Gly	Asp	His	Val					
			105					110						115						
gca	tgt	gcc	aga	act	ctg	ctg	gaa	gca	gga	gct	aat	gta	aat	gca	atc	859				
Ala	Cys	Ala	Arg	Thr	Leu	Leu	Glu	Ala	Gly	Ala	Asn	Val	Asn	Ala	Ile					
			120				125						130							
acg	ata	gat	ggc	gtg	act	ccg	tta	ttc	aac	gca	tgc	tcc	caa	ggc	agt	907				
Thr	Ile	Asp	Gly	Val	Thr	Pro	Leu	Phe	Asn	Ala	Cys	Ser	Gln	Gly	Ser					
		135					140					145								
cca	agc	tgt	gca	gag	ctg	ctt	ctg	gag	tat	ggg	gcc	aaa	gcc	cag	ctg	955				
Pro	Ser	Cys	Ala	Glu	Leu	Leu	Leu	Glu	Tyr	Gly	Ala	Lys	Ala	Gln	Leu					
	150					155					160									
gag	tca	tgt	ctt	cca	tcc	cca	acg	cat	gag	gcc	gcc	agt	aaa	ggg	cac	1003				
Glu	Ser	Cys	Leu	Pro	Ser	Pro	Thr	His	Glu	Ala	Ala	Ser	Lys	Gly	His					
165					170					175					180					
cat	gaa	tgt	ctt	gac	atc	ctg	ata	tcc	tgg	ggc	ata	gat	gtt	gac	caa	1051				
His	Glu	Cys	Leu	Asp	Ile	Leu	Ile	Ser	Trp	Gly	Ile	Asp	Val	Asp	Gln					
				185					190					195						
gaa	att	cct	cat	ttg	gga	act	cct	ctc	tat	gta	gct	tgt	atg	tca	cag	1099				
Glu	Ile	Pro	His	Leu	Gly	Thr	Pro	Leu	Tyr	Val	Ala	Cys	Met	Ser	Gln					
			200					205					210							
caa	ttc	cat	tgc	atc	tgg	aag	ctt	ctt	tat	gct	ggg	gct	gac	gta	cag	1147				
Gln	Phe	His	Cys	Ile	Trp	Lys	Leu	Leu	Tyr	Ala	Gly	Ala	Asp	Val	Gln					
		215					220					225								
aaa	ggc	aaa	tat	tgg	gat	act	cca	tta	cat	gct	gct	gct	caa	caa	tcc	1195				
Lys	Gly	Lys	Tyr	Trp	Asp	Thr	Pro	Leu	His	Ala	Ala	Ala	Gln	Gln	Ser					
	230					235					240									
agc	aca	gaa	att	gta	aac	tta	ctg	cta	gaa	ttt	gga	gca	gat	atc	aat	1243				
Ser	Thr	Glu	Ile	Val	Asn	Leu	Leu	Leu	Glu	Phe	Gly	Ala	Asp	Ile	Asn					

245	250	255	260	
gcc aaa aat aca gag ctt ctg cga cct ata gat gta gct acg tct agc				1291
Ala Lys Asn Thr Glu Leu Leu Arg Pro Ile Asp Val Ala Thr Ser Ser				
	265	270	275	
agt atg gtg gaa agg ata ttg ctt caa cat gaa gct acc cca agc tct				1339
Ser Met Val Glu Arg Ile Leu Leu Gln His Glu Ala Thr Pro Ser Ser				
	280	285	290	
ctt tac caa ctt tgc cga ctc tgt atc cga agc tac ata gga aaa cca				1387
Leu Tyr Gln Leu Cys Arg Leu Cys Ile Arg Ser Tyr Ile Gly Lys Pro				
	295	300	305	
aga ttg cac ctt atc cca caa ctc cag ctg cca acg tta ctg aag aat				1435
Arg Leu His Leu Ile Pro Gln Leu Gln Leu Pro Thr Leu Leu Lys Asn				
	310	315	320	
ttc tta cag tat cga taa aacagtaaag taattctaaa taccttgaaa				1483
Phe Leu Gln Tyr Arg				
325	330			
atcaaaattt ctatttcttt tgcttaagga atagttcata taaaaatatg ctaaagatag				1543
gataaaagtg agtgtgagat caccagggga agcagtaaaa tatcaatttt cattttaagt				1603
gtattagtagt tattgtttta tcatttatgc tgtctttgga ttatagcata tttataatat				1663
ctattttttg ttatttcaat ataccctctt gtttattcag gaaaaataa taaagtcact				1723
atcttatggt ttcttttgga ttggaaagaa ataaagctcc tagtaaagtt attcatatta				1783
tatttaaaaa ggtctatggt cttcaactaa tgggttaact ttttaatttt ggctattgta				1843
attggctaac ttgttactga tattgacata aaatatgaat atggcttagc aaagagcaac				1903
taatatgttt tctgaaacca ggatgatgct cattcaggaa atccactatg ttgaatgtcc				1963
actggtcatt agaattgaaca cttcataaag ataaattgcc cccctcccc ccacttacc				2023
ctggcaaagg actcacaagt cacttctctt ctccatagta catggacttg ttacagaaaa				2083
gtgataggat gaaagggtcaa tactgccctc tcaccctgag atatttagaa cattacacac				2143
cttctaaggc tatgttaaga ctgcctcctt ctccaggagta gccctgatg atctcttata				2203
aaggcgggta tatgcatatt actccagcct tcttccattt cagaaattgt caatagctag				2263
tgatacttgg acttaagaaa atgaactata tatctaatta acaaacaggc cctggatcat				2323
attccgggga tcatcacatc agtccaagga tataaaaaaa tggaagatta aaaactcaaa				2383
ttgtggagtt tttaaagagt ggggtgccact taacacattg gagaccttaa aaaatgtata				2443
tagagaactt aatcgactg tttctaattc tttttacaca aaacaagggc ttagagtata				2503

ctgtatttgc aagtaatagc aagagagtat tgtataaaag aagatattac taagctagtc 2563
 ctctgcaagg gtaaatacta tgtttattct cttcagcttt aactatactt gagtgcaata 2623
 aaaaacttaa taaaatttta tgccctgtttt gtaattcagg gcctattatg attttagaaa 2683
 taagataatc aacgtacaac tctttgtaaa acaaaagcaa ctttgggtga gttttataga 2743
 tttattaaat tttgataaag taaatatttt gaccactatt ttcaacaacc cagggcttat 2803
 tttcttctga cttcttggct aacttccaga ctatgggaaa ataaattatg attgtatgtg 2863
 acccccagat cactaagctt gcattttggc agaaagaaaa aaaatcagtt gtgatttaaa 2923
 acgtttgttg agattttcaa acatagaaat aaagctgtag catattcact aaacaatggt 2983
 ttatatgagg atagggacct ttaaatatgt gttttataga atatctcaca catttgtaat 3043
 ggataacgtg cattaaaatg gagttcttat tatagaataa cctgggttaa cagatattaa 3103
 agctgctttt gggtaaaaca ggtttccacc agtattttcc ctgagctaga ggagaagtct 3163
 tagattttta gagtaataca acccatctat cttattagtc ttgtgccaat aagagaagta 3223
 aatatcaaac attaaaatgg tcactttaca atttcagtat gtctgtacct tgtatcttag 3283
 cttgagtatg attgtcttat caaagtagtg ttctaacttt aatgtattgt atgtttgtta 3343
 ttatattaaa gtgtgattat tattcattgg aataaaaaaa aaaaaaaaaa aaaaa 3398

<210> 2

<211> 329

<212> PRT

<213> Homo sapiens

<400> 2

Met	Ser	Val	Leu	Glu	Glu	Asn	Arg	Pro	Phe	Ala	Gln	Gln	Leu	Ser	Asn
1				5					10					15	
Val	Tyr	Phe	Thr	Ile	Leu	Ser	Leu	Phe	Cys	Phe	Lys	Leu	Phe	Val	Lys
			20					25					30		
Ile	Ser	Leu	Ala	Ile	Leu	Ser	His	Phe	Tyr	Ile	Val	Lys	Gly	Asn	Arg
		35					40					45			
Lys	Glu	Ala	Ala	Arg	Ile	Ala	Ala	Glu	Phe	Tyr	Gly	Val	Thr	Gln	Gly
	50					55					60				
Gln	Gly	Ser	Trp	Ala	Asp	Arg	Ser	Pro	Leu	His	Glu	Ala	Ala	Ser	Gln
	65				70					75					80
Gly	Arg	Leu	Leu	Ala	Leu	Arg	Thr	Leu	Leu	Ser	Gln	Gly	Tyr	Asn	Val
			85					90						95	
Asn	Ala	Val	Thr	Leu	Asp	His	Val	Thr	Pro	Leu	His	Glu	Ala	Cys	Leu
			100					105					110		
Gly	Asp	His	Val	Ala	Cys	Ala	Arg	Thr	Leu	Leu	Glu	Ala	Gly	Ala	Asn
		115					120				125				
Val	Asn	Ala	Ile	Thr	Ile	Asp	Gly	Val	Thr	Pro	Leu	Phe	Asn	Ala	Cys
	130					135					140				
Ser	Gln	Gly	Ser	Pro	Ser	Cys	Ala	Glu	Leu	Leu	Leu	Glu	Tyr	Gly	Ala

145		150		155		160									
Lys	Ala	Gln	Leu	Glu	Ser	Cys	Leu	Pro	Ser	Pro	Thr	His	Glu	Ala	Ala
		165							170					175	
Ser	Lys	Gly	His	His	Glu	Cys	Leu	Asp	Ile	Leu	Ile	Ser	Trp	Gly	Ile
		180						185					190		
Asp	Val	Asp	Gln	Glu	Ile	Pro	His	Leu	Gly	Thr	Pro	Leu	Tyr	Val	Ala
		195					200					205			
Cys	Met	Ser	Gln	Gln	Phe	His	Cys	Ile	Trp	Lys	Leu	Leu	Tyr	Ala	Gly
	210					215				220					
Ala	Asp	Val	Gln	Lys	Gly	Lys	Tyr	Trp	Asp	Thr	Pro	Leu	His	Ala	Ala
225					230				235					240	
Ala	Gln	Gln	Ser	Ser	Thr	Glu	Ile	Val	Asn	Leu	Leu	Leu	Glu	Phe	Gly
		245						250					255		
Ala	Asp	Ile	Asn	Ala	Lys	Asn	Thr	Glu	Leu	Leu	Arg	Pro	Ile	Asp	Val
		260					265					270			
Ala	Thr	Ser	Ser	Ser	Met	Val	Glu	Arg	Ile	Leu	Leu	Gln	His	Glu	Ala
	275						280					285			
Thr	Pro	Ser	Ser	Leu	Tyr	Gln	Leu	Cys	Arg	Leu	Cys	Ile	Arg	Ser	Tyr
	290					295			300						
Ile	Gly	Lys	Pro	Arg	Leu	His	Leu	Ile	Pro	Gln	Leu	Gln	Leu	Pro	Thr
305				310				315					320		
Leu	Leu	Lys	Asn	Phe	Leu	Gln	Tyr	Arg							
			325												

<210> 3

<211> 3357

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (263)..(2191)

<223> Clone No. JT156897

<400> 3

gtttgttttc cccaacggga cgctctacat ccgcaacctc gcgccaagg acagcgggcg 60

ctatgagtgc gtggccgcca acctggtagg ctccgcgcgc aggacggtgc agctgaacgt 120

gcagcgtgca gcagccaacg cgcgcacacac gggcacctcc ccgcggagga cggacgtcag 180

gtacggagga accctcaagc tggactgcag cgcctcgggg gaccctggc cgcgcacacct 240

ctggaggctg ccgtccaaga gg atg atc gac gcg ctc ttc agt ttt gat agc 292

Met Ile Asp Ala Leu Phe Ser Phe Asp Ser

1

5

10

aga atc aag gtg ttt gcc aat ggg acc ctg gtg gtg aaa tca gtg acg 340

Arg Ile Lys Val Phe Ala Asn Gly Thr Leu Val Val Lys Ser Val Thr

15

20

25

gac aaa gat gcc gga gat tac ctg tgc gta gct cga aat aag gtt ggt 388

Asp Lys Asp Ala Gly Asp Tyr Leu Cys Val Ala Arg Asn Lys Val Gly

30

35

40

gat gac tac gtg gtg ctc aaa gtg gat gtg gtg atg aaa ccg gcc aag	436
Asp Asp Tyr Val Val Leu Lys Val Asp Val Val Met Lys Pro Ala Lys	
45 50 55	
att gaa cac aag gag gag aac gac cac aaa gtc ttc tac ggg ggt gac	484
Ile Glu His Lys Glu Glu Asn Asp His Lys Val Phe Tyr Gly Gly Asp	
60 65 70	
ctg aaa gtg gac tgt gtg gcc acc ggg ctt ccc aat ccc gag atc tcc	532
Leu Lys Val Asp Cys Val Ala Thr Gly Leu Pro Asn Pro Glu Ile Ser	
75 80 85 90	
tgg agc ctc cca gac ggg agt ctg gtg aac tcc ttc atg cag tcg gat	580
Trp Ser Leu Pro Asp Gly Ser Leu Val Asn Ser Phe Met Gln Ser Asp	
95 100 105	
gac agc ggt gga cgc acc aag cgc tat gtc gtc ttc aac aat ggg aca	628
Asp Ser Gly Gly Arg Thr Lys Arg Tyr Val Val Phe Asn Asn Gly Thr	
110 115 120	
ctc tac ttt aac gaa gtg ggg atg agg gag gaa gga gac tac acc tgc	676
Leu Tyr Phe Asn Glu Val Gly Met Arg Glu Glu Gly Asp Tyr Thr Cys	
125 130 135	
ttt gct gaa aat cag gtc ggg aag gac gag atg aga gtc aga gtc aag	724
Phe Ala Glu Asn Gln Val Gly Lys Asp Glu Met Arg Val Arg Val Lys	
140 145 150	
gtg gtg aca gcg ccc gcc acc atc cgg aac aag act tac ttg gcg gtt	772
Val Val Thr Ala Pro Ala Thr Ile Arg Asn Lys Thr Tyr Leu Ala Val	
155 160 165 170	
cag gtg ccc tat gga gac gtg gtc act gta gcc tgt gag gcc aaa gga	820
Gln Val Pro Tyr Gly Asp Val Val Thr Val Ala Cys Glu Ala Lys Gly	
175 180 185	
gaa ccc atg ccc aag gtg act tgg ttg tcc cca acc aac aag gtg atc	868
Glu Pro Met Pro Lys Val Thr Trp Leu Ser Pro Thr Asn Lys Val Ile	
190 195 200	
ccc acc tcc tct gag aag tat cag ata tac caa gat ggc act ctc ctt	916
Pro Thr Ser Ser Glu Lys Tyr Gln Ile Tyr Gln Asp Gly Thr Leu Leu	
205 210 215	
att cag aaa gcc cag cgt tct gac agc ggc aac tac acc tgc ttg gtc	964
Ile Gln Lys Ala Gln Arg Ser Asp Ser Gly Asn Tyr Thr Cys Leu Val	
220 225 230	
agg aac agc gcg gga gag gat agg aag acg gtg tgg att cac gtc aac	1012
Arg Asn Ser Ala Gly Glu Asp Arg Lys Thr Val Trp Ile His Val Asn	
235 240 245 250	
gtc cag cca ccc aag atc aac ggt aac ccc aac ccc atc acc acc gtg	1060
Val Gln Pro Pro Lys Ile Asn Gly Asn Pro Asn Pro Ile Thr Thr Val	
255 260 265	

cgg gag ata gca gcc ggg ggc agt cgg aaa ctg att gac tgc aaa gct	1108
Arg Glu Ile Ala Ala Gly Gly Ser Arg Lys Leu Ile Asp Cys Lys Ala	
270 275 280	
gaa ggc atc ccc acc ccg agg gtg tta tgg gct ttt ccc gag ggt gtg	1156
Glu Gly Ile Pro Thr Pro Arg Val Leu Trp Ala Phe Pro Glu Gly Val	
285 290 295	
gtt ctg cca gct cca tac tat gga aac cgg atc act gtc cat ggc aac	1204
Val Leu Pro Ala Pro Tyr Tyr Gly Asn Arg Ile Thr Val His Gly Asn	
300 305 310	
ggt tcc ctg gac atc agg agt ttg agg aag agc gac tcc gtc cag ctg	1252
Gly Ser Leu Asp Ile Arg Ser Leu Arg Lys Ser Asp Ser Val Gln Leu	
315 320 325 330	
gta tgc atg gca cgc aac gag gga ggg gag gcg agg ttg atc gtg cag	1300
Val Cys Met Ala Arg Asn Glu Gly Gly Glu Ala Arg Leu Ile Val Gln	
335 340 345	
ctc act gtc ctg gag ccc atg gag aaa ccc atc ttc cac gac ccg atc	1348
Leu Thr Val Leu Glu Pro Met Glu Lys Pro Ile Phe His Asp Pro Ile	
350 355 360	
agc gag aag atc acg gcc atg gcg ggc cac acc atc agc ctc aac tgc	1396
Ser Glu Lys Ile Thr Ala Met Ala Gly His Thr Ile Ser Leu Asn Cys	
365 370 375	
tct gcc gcg ggg acc ccg aca ccc agc ctg gtg tgg gtc ctt ccc aat	1444
Ser Ala Ala Gly Thr Pro Thr Pro Ser Leu Val Trp Val Leu Pro Asn	
380 385 390	
ggc acc gat ctg cag agt gga cag cag ctg cag cgc ttc tac cac aag	1492
Gly Thr Asp Leu Gln Ser Gly Gln Gln Leu Gln Arg Phe Tyr His Lys	
395 400 405 410	
gct gac ggc atg cta cac att agc ggt ctc tcc tcg gtg gac gct ggg	1540
Ala Asp Gly Met Leu His Ile Ser Gly Leu Ser Ser Val Asp Ala Gly	
415 420 425	
gcc tac cgc tgc gtg gcc cgc aat gcc gct ggc cac acg gag agg ctg	1588
Ala Tyr Arg Cys Val Ala Arg Asn Ala Ala Gly His Thr Glu Arg Leu	
430 435 440	
gtc tcc ctg aag gtg gga ctg aag cca gaa gca aac aag cag tat cat	1636
Val Ser Leu Lys Val Gly Leu Lys Pro Glu Ala Asn Lys Gln Tyr His	
445 450 455	
aac ctg gtc agc atc atc aat ggt gag acc ctg aag ctc ccc tgc acc	1684
Asn Leu Val Ser Ile Ile Asn Gly Glu Thr Leu Lys Leu Pro Cys Thr	
460 465 470	
cct ccc ggg gct ggg cag gga cgt ttc tcc tgg acg ctc ccc aat ggc	1732
Pro Pro Gly Ala Gly Gln Gly Arg Phe Ser Trp Thr Leu Pro Asn Gly	
475 480 485 490	

atg cat ctg gag ggc ccc caa acc ctg gga cgc gtt tct ctt ctg gac	1780
Met His Leu Glu Gly Pro Gln Thr Leu Gly Arg Val Ser Leu Leu Asp	
495 500 505	
aat ggc acc ctc acg gtt cgt gac gcc tcg gtg ttt gac agg ggt acc	1828
Asn Gly Thr Leu Thr Val Arg Asp Ala Ser Val Phe Asp Arg Gly Thr	
510 515 520	
tat gta tgc agg atg gag acg gag tac ggc cct tcg gtc acc agc atc	1876
Tyr Val Cys Arg Met Glu Thr Glu Tyr Gly Pro Ser Val Thr Ser Ile	
525 530 535	
ccc gtg att gtg atc gcc tat cct ccc cgg atc acc agc gag ccc acc	1924
Pro Val Ile Val Ile Ala Tyr Pro Pro Arg Ile Thr Ser Glu Pro Thr	
540 545 550	
ccg gtc atc tac acc cgg ccc ggg aac acc gtg aaa ctg aac tgc atg	1972
Pro Val Ile Tyr Thr Arg Pro Gly Asn Thr Val Lys Leu Asn Cys Met	
555 560 565 570	
gct atg ggg att ccc aaa gct gac atc acg tgg gag tta ccg gat aag	2020
Ala Met Gly Ile Pro Lys Ala Asp Ile Thr Trp Glu Leu Pro Asp Lys	
575 580 585	
tcg cat ctg aag gca ggg gtt cag gct cgt ctg tat gga aac aga ttt	2068
Ser His Leu Lys Ala Gly Val Gln Ala Arg Leu Tyr Gly Asn Arg Phe	
590 595 600	
ctt cac ccc cag gga tca ctg acc atc cag cat gcc aca cag aga gat	2116
Leu His Pro Gln Gly Ser Leu Thr Ile Gln His Ala Thr Gln Arg Asp	
605 610 615	
gcc ggc ttc tac aag tgc atg gca aaa aac att ctc ggc agt gac tcc	2164
Ala Gly Phe Tyr Lys Cys Met Ala Lys Asn Ile Leu Gly Ser Asp Ser	
620 625 630	
aaa aca act tac atc cac gtc ttc tga aatgtggatt ccagaatgat	2211
Lys Thr Thr Tyr Ile His Val Phe	
635 640	
tgcttaggaa ctgacaacaa agcgggggttt gtaagggaag ccagggttggg gaataggagc	2271
tcttaaataa tgtgtcacag tgcattggtgg cctctggtgg gtttcaagtt gaggttgatc	2331
ttgatctaca attggtggga aaaggaagca atgcagacac gagaaggagg gctcagcctt	2391
gctgagacac tttcttttgt gtttacatca tgccaggggc ttcattcagg gtgtctgtgc	2451
tctgactgca atttttcttt ttttgcaaat gccactcgac tgccttcata agcgtccata	2511
ggatatctga ggaacattca tcaaaaataa gccatagaca tgaacaacac ctactaccc	2571
cattgaagac gcatcaccta gttaacctgc tgcagttttt acatgataga ctttgttcca	2631
gattgacaag tcattcttca gttatttctt ctgtcacttc aaaactccag cttgccaat	2691

aaggatttag aaccagagtg actgatatat atatataatt taattcagag ttacatacat 2751
acagctacca ttttatatga aaaaagaaaa acatttcttc ctggaactca ctttttatat 2811
aatgttttat atatataatt tttcctttca aatcagacga tgagactaga aggagaaata 2871
ctttctgtct tattaaaatt aataaattat tgggtctttac aagacttgga tacattacag 2931
cagacatgga aatataattt taaaaaattt ctctccaacc tccttcaa at tcagtcacca 2991
ctgttatatt accttctcca ggaaccctcc agtggggaag gctgcatat tagatttcct 3051
tgtatgcaaa gtttttggtg aaagctgtgc tcagaggagg tgagaggaga ggaaggagaa 3111
aactgcatca taactttaca gaattgaatc tagagtcttc cccgaaaagc ccagaaactt 3171
ctctgcagta tctggcttgt ccatctggtc taaggtggct gcttcttccc cagccatgag 3231
tcagtttgtg cccatgaata atacacgacc tggtatttcc atgactgctt tactgtattt 3291
ttaaggtcaa tatactgtac atttgataat aaaataatat tctcccaaaa aaaaaaaaaa 3351
aaaaaa 3357

<210> 4

<211> 642

<212> PRT

<213> Homo sapiens

<400> 4

Met	Ile	Asp	Ala	Leu	Phe	Ser	Phe	Asp	Ser	Arg	Ile	Lys	Val	Phe	Ala	1	5	10	15
Asn	Gly	Thr	Leu	Val	Val	Lys	Ser	Val	Thr	Asp	Lys	Asp	Ala	Gly	Asp	20	25	30	
Tyr	Leu	Cys	Val	Ala	Arg	Asn	Lys	Val	Gly	Asp	Asp	Tyr	Val	Val	Leu	35	40	45	
Lys	Val	Asp	Val	Val	Met	Lys	Pro	Ala	Lys	Ile	Glu	His	Lys	Glu	Glu	50	55	60	
Asn	Asp	His	Lys	Val	Phe	Tyr	Gly	Gly	Asp	Leu	Lys	Val	Asp	Cys	Val	65	70	75	80
Ala	Thr	Gly	Leu	Pro	Asn	Pro	Glu	Ile	Ser	Trp	Ser	Leu	Pro	Asp	Gly	85	90	95	
Ser	Leu	Val	Asn	Ser	Phe	Met	Gln	Ser	Asp	Asp	Ser	Gly	Gly	Arg	Thr	100	105	110	
Lys	Arg	Tyr	Val	Val	Phe	Asn	Asn	Gly	Thr	Leu	Tyr	Phe	Asn	Glu	Val	115	120	125	
Gly	Met	Arg	Glu	Glu	Gly	Asp	Tyr	Thr	Cys	Phe	Ala	Glu	Asn	Gln	Val	130	135	140	
Gly	Lys	Asp	Glu	Met	Arg	Val	Arg	Val	Lys	Val	Val	Thr	Ala	Pro	Ala	145	150	155	160
Thr	Ile	Arg	Asn	Lys	Thr	Tyr	Leu	Ala	Val	Gln	Val	Pro	Tyr	Gly	Asp	165	170	175	
Val	Val	Thr	Val	Ala	Cys	Glu	Ala	Lys	Gly	Glu	Pro	Met	Pro	Lys	Val	180	185	190	
Thr	Trp	Leu	Ser	Pro	Thr	Asn	Lys	Val	Ile	Pro	Thr	Ser	Ser	Glu	Lys				

10

